

**MOLECULAR INVESTIGATION OF *RAD51* AND  
*DMC1* HOMOEOLGOUS GENES OF HEXAPLOID  
WHEAT (*Triticum aestivum* L.)**

**BY**

**UPENDRA KUMAR DEVISETTY  
MSc (Ag).,**

UNIVERSITY OF NOTTINGHAM <  
JAMES CAMERON-GIFFORD LIBRARY

**Thesis submitted to the University of Nottingham for  
degree of Doctor of Philosophy**

**November 2009**

# Table of Contents

<b>Abstract</b> .....	I
<b>Acknowledgements</b> .....	III
<b>Abbreviations</b> .....	IV
<b>Chapter 1: General Introduction</b> .....	1
1.1 General information of wheat ( <i>Triticum aestivum</i> L.).....	1
1.2 History of wheat.....	2
1.3 Polyploidy in plants.....	2
1.4 Origin of A, B and D genomes of bread wheat.....	5
1.5 Different growth stages of wheat.....	8
1.6 Reproductive Biology of wheat.....	10
1.7 General information of <i>Arabidopsis thaliana</i> .....	11
1.7.1 Arabidopsis as a model plant.....	11
1.7.2 Growth stages of Arabidopsis.....	12
1.8 Meiosis.....	15
1.8.1 An overview.....	15
1.8.2 The anther and stages of meiosis in wheat.....	16
1.8.3 Chromosome pairing during meiosis in eukaryotes.....	18
1.8.3.1 Chromosome pairing in <i>Triticum aestivum</i> .....	19
1.9 Homologous Recombination.....	21
1.9.1 An overview.....	21
1.9.2 HR at meiosis.....	21
1.9.3 Models of HR.....	22
1.9.3.1 The Holliday model and Meselson-Radding models.....	22
1.9.3.2 The Double-Strand Break Repair (DSBR) model.....	25
1.9.4 HR in <i>E. coli</i> .....	27
1.9.5 HR during meiosis in <i>Saccharomyces cerevisiae</i> .....	30
1.9.6 HR in plants.....	31
1.10 Recombinases.....	33
1.10.1 <i>E. coli</i> RecA recombinase.....	33
1.10.2 Rad51 and Dmc 1 recombinases.....	34
1.11 <i>RAD51</i> gene.....	36



1.11.1 Structure of the Rad51 protein.....	38
1.12 <i>DMC1</i> gene.....	39
1.12.1 Dmc1 protein structure.....	41
1.13 <i>RAD51</i> paralogues.....	41
1.14 Phylogenetic analysis of Rad51 and Dmc1 proteins.....	44
1.15 Evolutionary history of RecA/RAD51 gene family.....	45
1.16 Other meiotic recombination genes.....	47
1.17 Meiotic Recombination and wheat breeding.....	47
1.18 HR and gene targeting.....	52
1.19 General project aims.....	53
<b>Chapter 2: Materials and Methods.....</b>	<b>57</b>
2.1 Materials.....	57
2.1.1 List of chemicals used.....	57
2.1.2 List of Equipment & Apparatus used.....	59
2.1.3 Oligonucleotides.....	60
2.1.4 Computer analysis software used.....	60
2.1.5 Plant materials used and growth conditions.....	62
2.1.6 Bacterial strains used.....	63
2.1.7 Cloning vectors used.....	63
2.2 Methods.....	64
2.2.1 Bacterial culture growth media.....	64
2.2.1.1 LB (Luria-Bertani) liquid medium.....	64
2.2.1.2 LB Agar.....	64
2.2.1.3 <i>E. coli</i> media supplements.....	64
2.2.1.4 <i>Agrobacterium tumefaciens</i> media supplements.....	64
2.2.2 Plant culture media.....	65
2.2.2.1 1/2 Murashige and Skoog (MS) solid media.....	65
2.2.2.2 MS media supplements.....	65
2.2.3 Antibiotic Stocks.....	65
2.2.3.1 Ampicillin (50mg/ml).....	65
2.2.3.2 Kanamycin (10mg/ml).....	65
2.2.3.3 Rifampicin (4mg/ml).....	65

2.2.4	<i>E. coli</i> Competent Cells preparation buffers.....	65
2.2.4.1	SOB.....	65
2.2.4.2	SOC.....	66
2.2.4.3	TB Buffer.....	66
2.2.4.4	10mM Pipes.....	66
2.2.4.5	TBE Buffer (5X).....	66
2.2.4.6	0.5M EDTA (pH 8.0).....	66
2.2.4.7	TE Buffer (1X).....	66
2.2.4.8	1M Tris HCl (pH 8.0).....	67
2.2.5	General reagents preparation.....	67
2.2.5.1	X-Gal (50mg/ml).....	67
2.2.5.2	Isopropylthio-b-D-galactoside (100 mM).....	67
2.2.5.3	Gel Loading Buffer (6x).....	67
2.2.6	Colored Stains.....	67
2.2.6.1	Alexander Stain.....	67
2.2.6.2	1% Aceto-Orcein.....	68
2.2.7	DNA methods.....	68
2.2.7.1	Plant genomic DNA extraction.....	68
2.2.7.1.1	Quantification of genomic DNA.....	71
2.2.7.2	Plasmid DNA extraction.....	71
2.2.7.2.1	Plasmid miniprep extraction.....	71
2.2.7.2.2	Plasmid Midiprep extraction.....	72
2.2.7.3	Isolation of DNA fragments from agarose gels.....	73
2.2.7.4	Purification of PCR products.....	73
2.2.7.5	DNA sequencing.....	74
2.2.8	RNA methods.....	75
2.2.8.1	Total RNA extraction.....	75
2.2.8.2	Quantification of RNA.....	76
2.2.9	First strand cDNA synthesis.....	76
2.2.9.1	DNase treatment of RNA.....	76
2.2.9.2	Reverse Transcription.....	76
2.2.10	PCR (Polymerase Chain Reaction) methods.....	77
2.2.10.1	Gradient PCR method.....	77

2.2.10.2	General and Phusion PCR.....	77
2.2.10.3	Long Amp Taq DNA PCR.....	78
2.2.10.4	Colony PCR.....	78
2.2.10.5	RT-PCR.....	79
2.2.10.6	Quantitative real-time PCR (Q-PCR).....	79
2.2.11	Agarose Gel Electrophoresis.....	79
2.2.12	DNA cloning methods.....	80
2.2.12.1	Restriction digestions for cloning.....	80
2.2.12.2	Ligation into the pGEM T-Easy vector system.....	81
2.2.12.3	Ligation into the pGVT17 vector system.....	82
2.2.13	<i>E. coli</i> methods.....	83
2.2.13.1	<i>E. coli</i> competent cells preparation.....	83
2.2.13.2	<i>E. coli</i> transformation using heat shock.....	84
2.2.13.3	X-gal + IPTG (Blue / white) selection of <i>E. coli</i> transformants.....	84
2.2.13.4	Glycerol Stock preparation.....	84
2.2.14	<i>Agrobacterium tumefaciens</i> methods.....	84
2.2.14.1	Preparation of <i>Agrobacterium</i> electrocompetent Cells.....	84
2.2.14.2	Transformation of <i>Agrobacterium</i> by Electroporation.....	85
2.2.15	Wheat methods.....	84
2.2.15.1	<i>In vitro</i> germination of wheat seeds.....	85
2.2.15.2	Mutagenesis methods.....	86
2.2.15.2.1	Fast-neutron mutagenesis.....	86
2.2.15.2.2	Gamma ray mutagenesis.....	87
2.2.15.3	Collection and meiotic staging of wheat anthers.....	87
2.2.15.4	Rapid Amplification of cDNA ends (RACE).....	88
2.2.15.4.1	3'RACE.....	88
2.2.15.4.2	5'RACE.....	89
2.2.15.5	Alexander staining.....	90
2.2.15.6	Cytogenetic analysis.....	91
2.2.15.7	Fertility assessment.....	91

2.2.16 Arabidopsis methods.....	91
2.2.16.1 <i>In vitro</i> germination of seeds.....	91
2.2.16.2 Transformation of Arabidopsis by floral dipping.....	92
2.2.16.2.1 Screening of Arabidopsis transformants.....	92
2.2.16.3 Genotypic analysis of the Arabidopsis <i>rad51</i> and <i>dmc1</i> T-DNA mutants.....	93
2.2.16.3.1 Genotypic analysis of the Arabidopsis <i>rad51</i> T-DNA lines.....	93
2.2.16.3.2 Genotypic analysis of the Arabidopsis <i>dmc1</i> T-DNA lines.....	94
2.2.16.3.3 Ligation of the CaMV35S:: <i>TaRAD51</i> (D)::± GFP and CaMV35S:: <i>TaDMC1</i> (D)::± GFP constructs for transformation into Arabidopsis <i>rad51</i> and <i>dmc1</i> T-DNA mutants, respectively.....	95
2.2.17 Micro array methods.....	98
2.2.17.1 Microarray data access.....	98
2.2.17.2 Probe set identification .....	98
2.2.17.3 BLAST for probe specificity.....	98
2.2.17.4 GeneSpring Software.....	99

<b>Chapter 3: Microarray Expression profiling of <i>RAD51</i> and <i>DMC1</i> in Arabidopsis, rice and wheat.....</b>	<b>100</b>
3.1 Introduction.....	100
3.2 Results.....	108
3.2.1 Identification of the probe sets corresponding to the <i>RAD51</i> and <i>DMC1</i> genes of Arabidopsis, rice and wheat..	108
3.2.2 Gene specificity of probe sequences of <i>RAD51</i> and <i>DMC1</i> in Arabidopsis, rice and wheat.....	109
3.2.3 Transcriptional expression profiling of the <i>RAD51</i> gene in Arabidopsis, rice and wheat using microarray databases...	118
3.2.4 Transcriptional expression profiling of the <i>DMC1</i> genes in Arabidopsis, rice and wheat using microarray databases...	125

3.3 Discussion.....	135
3.3.1 <i>RAD51</i> gene expression in Arabidopsis, rice and wheat...	135
3.3.2 <i>DMC1</i> gene expression in Arabidopsis, rice and Wheat...	136

## **Chapter 4: Cloning and molecular characterization of *RAD51* and**

<b><i>DMC1</i> homoeologous genes in <i>Triticum aestivum</i> L.....</b>	<b>140</b>
4.1 Introduction.....	140
4.2 Results.....	143
4.2.1 Isolation and molecular cloning of wheat <i>RAD51</i> cDNA sequence.....	143
4.2.2 Isolation and molecular cloning of the wheat <i>DMC1</i> cDNA sequence.....	144
4.2.3 Molecular cloning and chromosome location of <i>TaRAD51</i> and <i>TaDMC1</i> cDNA homoeologous genes.....	144
4.2.3.1 Development of Genome Specific Primer sets (GSPs) for the <i>TaRAD51</i> and <i>TaDMC1</i> genes.....	144
4.2.3.2 Isolation of the full-length <i>TaRAD51</i> and <i>TaDMC1</i> cDNA homoeologous genes.....	150
4.2.4 Analysis of amino acid sequence homology of full- length <i>TaRAD51</i> and <i>TaDMC1</i> homoeologous genes.....	152
4.2.5 Phylogenetic analysis of the <i>TaRAD51</i> and <i>TaDMC1</i> homoeologue genes.....	159
4.2.6 Comparative protein analysis of the <i>TaRAD51</i> and <i>TaDMC1</i> homoeologous genes.....	163
4.2.7 Structural conservation of the protein sequences of the <i>TaRAD51</i> and <i>TaDMC1</i> cDNA homoeologous genes.	166
4.2.8 Homoeolog-specific expression patterns of <i>TaRAD51</i> and <i>TaDMC1</i> genes.....	167
4.3 Discussion.....	170
4.3.1 The <i>RAD51</i> gene in hexaploid wheat.....	170
4.3.2 The <i>DMC1</i> gene in hexaploid wheat.....	173

**Chapter 5: Functional characterization of the *RAD51* and *DMC1***

**genes in *Triticum aestivum* L.....177**

5.1 Introduction.....177

5.2 Results.....181

5.2.1 Design of exon-anchored primers for the *TaRAD51*  
and *TaDMC1* genes.....181

5.2.2 Genome-Specific Primer design for the *TaRAD51*  
and *TaDMC1* genes.....182

5.2.2.1 *TaRAD51* Genome-Specific Primers.....182

5.2.2.2 *TaDMC1* Genome-Specific Primers.....183

5.2.3 Forward and Reverse Genetic screening of the  
Highbury deletion lines.....185

5.2.3.1 Estimation of mutation frequency in the  
Highbury deletion lines.....185

5.2.3.2 Forward genetic screening of Highbury deletion  
lines.....186

5.2.3.2.1 Genetic analysis of the identified  
Highbury deletion lines identified by  
Forward Genetic screening.....188

5.2.4 Forward and Reverse Genetic screening of the  
Paragon deletion lines.....189

5.2.4.1 Estimating mutation frequencies in the  
Paragon population.....189

5.2.4.2 Reverse Genetics screening of the Paragon  
deletion population.....189

5.2.4.2.1 Screening for deletions with *TaRAD51*  
Genome-Specific Primers.....190

5.2.4.2.2 Screening for deletions with *TaDMC1*  
Genome-Specific Primers.....191

5.2.4.2.3 Forward genetic screening of M<sub>3</sub> progeny  
of the confirmed *TaRAD51* and  
*TaDMC1* homoeologous gene mutants.....193

5.2.4.2.3.1 Confirmation of the transmission of the deletion and Forward Genetic screening of the <i>TaRAD51</i> homoeologous gene mutants.....	193
5.2.4.2.3.2 Confirmation of the transmission of the deletion and Forward Genetic screening of the <i>TaDMC1</i> homoeologous gene mutants...	194
5.2.4.2.4 <i>TaRAD51</i> and <i>TaDMC1</i> gene expression analysis in the identified Paragon deletion lines and unmutated Paragon.....	196
5.2.4.2.4.1 <i>TaRAD51</i> expression analysis.....	196
5.2.4.2.4.2 <i>TaDMC1</i> expression analysis.....	196
5.2.4.2.5 Pollen viability assay in the identified Paragon deletion lines of <i>TaRAD51</i> and <i>TaDMC1</i> homoeologous genes and unmutated Paragon.....	198
5.2.4.2.6 Meiotic assessment of identified Paragon deletion lines of the <i>TaRAD51</i> and <i>TaDMC1</i> homoeologous genes and comparison with unmutated Paragon.....	200
5.2.4.2.7 Fertility assessment of identified Paragon deletion lines of <i>TaRAD51</i> and <i>TaDMC1</i> homoeologous genes and unmutated Paragon.....	202
5.3 Discussion.....	204

## **Chapter 6: Functional characterization of *TaRAD51* and *TaDMC1* genes in *Arabidopsis thaliana***

6.1 Introduction.....	210
6.2 Results.....	215
6.2.1 Comparative amino acid sequence analysis between the <i>AtRAD51</i> and the <i>TaRAD51</i> homoeologous genes.....	215

6.2.2	Comparative amino acid sequence analysis between the <i>AtDMC1</i> and <i>TaDMC1</i> homoeologous genes.....	216
6.2.3	Comparative sequence analysis of the <i>TaRAD51(D)</i> and <i>AtRAD51</i> amino acids.....	218
6.2.4	Comparative sequence analysis of the <i>TaDMC1(D)</i> and <i>AtDMC1</i> amino acids.....	218
6.2.5	Structural conservation of the protein sequences of the <i>TaRAD51</i> and <i>AtRAD51</i> .....	221
6.2.6	Structural conservation of the protein sequences of the <i>TaDMC1</i> and <i>AtDMC1</i> .....	221
6.2.7	Molecular and phenotypic characterization of the <i>Atrad51</i> and <i>Atdmc1</i> T-DNA insertional lines.....	222
6.2.7.1	Molecular and phenotypic characterization of Arabidopsis <i>rad51</i> T-DNA insertional mutants...	223
6.2.7.2	Molecular and phenotypic characterization of Arabidopsis <i>dmc1</i> T-DNA insertion mutants.....	225
6.2.7.3	Expression analysis of the <i>AtRAD51</i> and <i>AtDMC1</i> T-DNA insertional mutants.....	227
6.2.7.4	Genetic transformation of <i>CaMV35S::TaRAD51(D)::±GFP</i> complementation constructs into the <i>Atrad51</i> T-DNA insertional mutant.....	228
6.2.7.4.1	Genetic transformation of the <i>CaMV35S::TaRAD51(D)</i> complementation construct into the <i>Atrad51</i> T-DNA insertional mutant.....	228
6.2.7.4.2	Genetic transformation of the <i>CaMV35S::TaRAD51(D)::GFP</i> construct into the <i>Atrad51</i> T-DNA insertional mutant line.....	231
6.2.7.5	Genetic transformation of <i>CaMV35S::TaDMC1(D)::±GFP</i> complementation constructs into the <i>AtDMC1</i> T-DNA insertional mutation line.....	233



6.2.7.5.1 Genetic transformation of <i>CaMV35S::TaDMC1(D)</i> construct into the <i>AtDMC1</i> T-DNA insertional mutant line...	234
6.2.7.5.2 Genetic transformation of the <i>CaMV35S::TaDMC1(D)::GFP</i> construct into the <i>AtDMC1</i> T-DNA insertional mutant line...	236
6.2.7.6 Confirmation of <i>TaRAD51(D)</i> expression in <i>Atrad51</i> homozygous transformed plants.....	242
6.2.7.7 Complementation of the <i>Atrad51</i> insertional mutation line with the over-expressing <i>CaMV35S::TaRAD51(D)::±GFP</i> construct.....	243
6.2.7.7.1 Complementation with the <i>CaMV35S::TaRAD51(D)</i> construct.....	243
6.2.7.7.2 Complementation with the <i>CaMV35S::TaRAD51(D)::GFP</i> construct.....	243
6.2.8 Effect of over-expression of the <i>TaRAD51</i> gene on genetic distances and recombination frequencies.....	256
6.3 Discussion.....	261
6.3.1 Wheat <i>RAD51</i> functional studies in Arabidopsis T-DNA lines.....	261
6.3.2 Wheat <i>DMC1</i> functional studies in Arabidopsis T-DNA lines.....	265
<b>Chapter 7: General discussion.....</b>	<b>268</b>
7.1 Future research.....	283
7.1.1 Isolation of full-length genomic clones of <i>TaRAD51</i> and <i>TaDMC1</i> homoeologous genes.....	283
7.1.2 Screening for the isolation of <i>TaRAD51 (A), (D)</i> and <i>TaDMC1 (B)</i> homoeologous gene mutants by Reverse Genetic analysis.....	283
7.1.3 Over-expression studies of the <i>TaDMC1</i> gene in <i>Atdmc1</i> homozygous mutants for possible complementation.....	283

7.1.4 Use of over-expressing constructs with gene specific ( <i>RAD51/DMC1</i> ) promoters.....	285
7.1.5 Studying the recombination frequency in over-expressing lines of <i>TaDMC1</i> .....	285
<b>References.....</b>	<b>286</b>

## Abstract

Meiotic recombination in eukaryotes requires two orthologues of the *E. coli* RecA proteins, Rad51 and Dmc1. Both genes play an important role in the binding of single strand DNA, homology search, strand invasion and strand exchange resulting in Holliday junctions which are resolved into crossovers or non-crossovers events. Even though both genes are well characterized in a variety of organisms including plants, very little information is available from hexaploid wheat. In most diploid plant species, deletion of either the *RAD51* or *DMC1* orthologues leads to sterility but wheat being a polyploid, offers a unique opportunity to examine the effects of the deletion of specific homoeologue, while maintaining a degree of fertility. The transcript expression profiling of *RAD51* and *DMC1* genes in Arabidopsis, rice and wheat using available microarray databases indicated higher levels of expression in mitotically and meiotically active tissues compared to other tissues. However, the possible function of the *DMC1* gene in mitotic-active tissues needs to be investigated further. Previously cDNA sequences of *TaRAD51* and *TaDMC1* of hexaploid wheat were cloned and reported. In this study, it has been demonstrated that the reported *TaRAD51A1* and *TaRAD51A2* cDNA sequences are (*D*) and (*A*) homoeologues of *TaRAD51* respectively and *TaDMC1* cDNA sequence is (*D*) homoeologue of the *TaDMC1*. This study also found that the amino acid sequences and evolutionary relationships of *RAD51* and *DMC1* cDNA homoeologues are highly conserved across eukaryotes. Functional characterization of *TaRAD51* and *TaDMC1* gene homoeologues was undertaken *in planta* using Forward Genetics, Reverse Genetics and Complementation methods. Forward and Reverse Genetic screening of a subset of a Highbury mutant population could not identify any mutants that have deletions in *TaRAD51* and *TaDMC1* genes. However, Reverse Genetics screening of Paragon mutant population identified mutant lines that tested as having deletions for all the three homoeologues of *TaRAD51* and *TaDMC1*. However, most likely due to high mutational load and a deleterious phenotype,

only a few mutant lines survived. Phenotypic and cytogenetic analysis indicated the probable functional redundancy of *TaRAD51(B)* homoeologue in meiosis, although the unknown size of the deletion and limited phenotype makes it impossible to completely certain of this. The single mutants for *TaDMC1(B)* and *(D)* indicated a reduction in pollen viability and ear fertility compared to wild-type. The cytological examination of these mutants indicated low levels of abnormal diakinesis, resulting in the formation of dyads. However, the single mutants were still able to produce normal tetrads. This suggests that there is a possible dosage effect of these homoeologues in hexaploid wheat. Unless deletion lines for the *(A)* and *(D)* homoeologues of *TaRAD51* and *(B)* homoeologue of *TaDMC1* can be recovered and characterized the above assumptions will remain inconclusive. The results of complementation assays using over-expressing *CaMV35S::TaRAD51(D)±GFP* constructs demonstrated a very low (~14% and ~2%, respectively, with +GFP and –GFP constructs) functional complementation in terms of seed set compared to 0% in homozygous *Atrad51* mutants. One explanation of these results is that the wheat genes are not complete functional orthologues for the inactivated *Arabidopsis* genes. The functional complementation experiments could not be performed for *TaDMC1* gene because of time limitation, although the transformants were produced in *AtDMC1/atdmc1* background. Finally, overexpression of the *TaRAD51* gene suggests 2-fold increase in genetic distances in *Arabidopsis* using *CaMV35S::TaRAD51(D)* construct. This was done by crossing the appropriate transformant with fluorescent tetrad lines. However the results need to be confirmed by a large scale analysis.

## **Acknowledgement**

I gratefully avail this opportunity to express my humble and deep sense of gratitude and regards to Dr. Sean Mayes for his precious guidance, invincible thoughts, splendid cooperation and endurance. Whatever I could do in this study might not have been possible without his constant encouragement, moral boosting, untiring supervision and constructive criticism. I express my deep gratitude to Dr. Katie Mayes and Dr. Tineke Sonneveld for their keen interest, congenial discussion and valuable technical guidance.

I record my sincere thanks to Dr. Simon Griffiths and Simon Orford from John Innes Centre, Dr. Marie-Pascale Doutriaux from Paris-Sud University and Dr. Copenhaver from the University of North Carolina for providing valuable materials for this research.

Thanks are also extended to Prof Mike Holdsworth and Dr. Tim Robbins and other members of the division for their sincere help and co-operation.

A deep sense of gratitude is owned to Fiona Wilkinson, Julietta Marquez, Sue Golds, Emma Hooley, Chris Mills, Sheila Northover, Sue Flint, Mark Meacham, Ros Beacham for their kind help and necessary facilities during the course of my study.

I am highly thankful to Sony, Hongying Li, Neil Graham, Dr. Shravani Basu, Dr. Chanate Malumpong, Stephen Hayes, Dr George, Dr Kingsley, Dr Gracia, Ozie, Endah, Rafal, Daniel for their unmatched help and support from the start till the preparation of this thesis.

A special thanks to my family and friends for all their support and constant encouragement throughout the course of my PhD.

Last but not least Full Tuition Fee Research Scholarship provided by The University of Nottingham is duly acknowledged.

## Abbreviations

2/3D	2/3 Dimensional
ACC1	Acetyl- CoA Carboxylase 1
AG	Affymetrix AtGenome
AGI	Arabidopsis Genome Initiative
ASY	Asynaptic
ATP	Adenosine 5'-Triphosphate
BCDX2	Rad51B, Rad51C, Rad51D and Xrcc2 complex
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
CaMV	Cauliflower Mosaic Virus
CDK	Cyclin Dependant Kinases
cDNA	complementary Deoxyribonucleic acid
CFP	Cyan Fluorescent Protein
CHI	Crossover Hotspot Initiator
CO	cross-over
CX3	Rad51C and Xrcc3 complex
DArT	Diversity Array Technology
DH	Double Haploid
DMC1	Disruption of Meiotic Control 1
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide Tri Phosphate
ds	double strand
DSB	Double-Strand Break
DSBR	Double-Strand Break Repair
dsDNA	double-strand DNA
EBI	European Bioinformatics Institute
ECFP	Enhanced CFP
EDTA	Ethylenediaminetetraacetic acid
EF-G	Elongation Factor G
eFP	electronic Fluorescent Pictograph

EST	Expressed Sequence Tag
EYFP	Enhanced YFP
FISH	Fluorescent In Situ Hybridization
FL-cDNA	Full-length cDNA
FST	Flanking sequence tags
FTL	Fluorescent Tetrad Line
GC	Gene Conversion
gDNA	genomic DNA
GEO	Gene Expression Omnibus
GFP	Green Fluorescent Protein
Glu	Glutamic acid
Gly	Glycine
GSP	Genome Specific Primer
GT	Gene Targeting
HJs	Holliday Junctions
HOP2	Homologous Pairing 2
HR	Homologous Recombination
I <sub>1</sub>	Interval 1
I <sub>3</sub>	Interval 3
JM	Joint Molecule
JMs	Joint Molecule recombination intermediates
LIM15	Lily messages Induced in Meiosis 15
MEGA	Molecular Evolutionary Genomic Analysis
MEI4	Meiosis specific 4
MEI5	Meiosis specific 5
Met	Methionine
MM	Mismatch
MMC	MitomycinC
MMS	Methyl methanesulfonate
MND1	Meiotic Nuclear Divisions 1
MRE11	Meiotic Recombination 11
MRN	Mre11-Rad51-Nbs1
mRNA	messenger Ribonucleic acid
MRX	Mre11-Rad50-Xrs2

MS	Murashige and Skoog
MSH	MutS Homolog
NASC	Nottingham Arabidopsis Stock Centre
NBS1	Nijmegen Breakage Syndrome 1
NCBI	National Center for Biotechnology Information
NHEJ	Non-Homologous End Joining
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
Pgk1	3-Phosphoglycerate kinase
Ph1	Pairing homoeologous 1
Phs1	Poor homologous synapsis 1
pLEXdb	Plant Expression Database
PM	Perfect Match
PTD	Parting Dancers
Q-PCR	Quatitative Polymerase Chain Reaction
RACE	Rapid Amplification of cDNA Ends
RAD50	Radiation sensitive 50
RAD51	Radiation sensitive 51
RAD54	Radiation sensitive 54
RecA	Recombination Deficient A
RED	Rice Expression Database
RF	Recombination Frequency
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic acid
RNAi	RNA interference
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SC	Synaptonemal Complex
SDSA	Synthesis-Dependent Strand-Annealing
SIFT	Sorting Intolerant From Tolerant
SPO11	Sporulation 11
ss	single-strand
SSCP	Single Strand Conformation Polymorphism
TA	Triticum Aestivum
TAIR	The Arabidopsis Information Resource



tBLASTX	Translated Basic Local Alignment Search Tool
T-DNA	Transfer-Deoxyribo Nucleic Acid
TIGR	The Institute for Genomic Research
URS1	Upstream Repressor Sequence 1
UV	Ultraviolet
XRCC2	X-Ray Cross Complementing 2
XRCC3	X-Ray Cross Complementing 3
XRS2	X-Ray Sensitive 2
XRS4	X-Ray Sensitive 4
YFP	Yellow Fluorescent Protein
ZIP1	Zipper 1

# CHAPTER 1: GENERAL INTRODUCTION

## 1.1 General information of wheat (*Triticum aestivum* L.)

Common wheat or Bread wheat (*Triticum aestivum* L.) is one of the world's most important food crops along with rice and maize. It is a staple food crop for many countries in the world. The wheat is an annual plant and for its early growth thrives well in cooler conditions. Wheat belongs to the tribe Triticeae, which includes 15 genera and over 300 species including wheat, rye and barley (Sakamoto, 1973). Bread wheat is a part of the grass family Poaceae that includes other major cereal crops such as oats, maize and rice. Among the more ancient and now less frequently cultivated species are einkorn (*T. monococcum*), emmer (*T. dicoccum*) and spelt (*T. spelta*) wheats. Modern wheat varieties are either classified as winter wheats (those which require vernalization requirements for few weeks before planting) or spring wheats (those that does not require vernalization requirements). The flour from hard wheats (varieties evolved for the most part from *T. aestivum*) contains a high percentage of gluten and is generally used for making bread and cakes. White- and soft-wheat varieties are paler and have starchy kernels and their flour is preferred for piecrust, biscuits, and breakfast foods. Wheat is also used in the manufacture of whiskey and beer. The grain, the bran (the residue from milling) and the vegetative plant parts make valuable livestock feed. Wheat was once used as the principal source of starch for sizing paper and cloth.

Apart from the humankind's most important crop, bread wheat is also an excellent model for the analysis of basic mechanisms of genome evolution and speciation as well as for genetic approaches for crop improvement. Wheat is a model system for polyploid species because of relative ease in chromosome identification and manipulations, and the presence of a wealth of genetic and cytogenetic stocks. Wheat is also best suited for understanding some of the fundamental biological processes such as *Phl* Sears (1976).

## 1.2 History of wheat

Wheat evolved from wild grasses, probably somewhere in the Near East but the precise origin of wheat is not known even today. The Fertile Crescent – “a region with rich soils in the upper reaches of the Tigris-Euphrates drainage basin” (Briggle and Curtis, 1987) is the most likely place for the origin of wheat. The cultivation of ‘naked’ wheats began between the late fifth and early fourth millennium B.C in the southern Caucasus region in neolithic settlements. Archaeological evidence also shows that naked wheat was found at several sites in the Crimea (1000-900 B.C.) which matches the archaeological finding of wheat in Israel at the same period (Korber-Grohne, 1988).

## 1.3 Polyploidy in plants

The term polyploidy was coined by Winkler (1916). Polyploidy is a wide spread phenomenon occurring in many plants (Ohno, 1970 and Wolfe, 2001) but predominantly among higher plants (angiosperms) including many important agricultural crops such as wheat, oat, tobacco, potato, banana, sugarcane, cotton, and certain *Brassica* species (Leitch and Bennett, 1997). Polyploidy has had a major effect on the evolution of flowering plants and has made a considerable impact on plant species diversity by the way of providing raw material for evolution and relaxing purifying selection on the duplicated genes in the polyploid (Wendel and Doyle, 2005). Among angiosperms, 50-70% of the species are polyploids (Stebbins, 1971; Soltis and Soltis, 1993 and Masterson, 1994). Molecular analyses such as the use of RFLP markers and more recently genome sequencing have shown that even presumed diploids such as maize and *Arabidopsis* are in fact ancient polyploids (Wendel, 2000). The major advantage of polyploids over their diploid progenitors is their potential to adapt to a wide variety of habitats and survive successfully in unstable climates (Ehrendorfer, 1980; Levin, 1983; Novak *et al.*, 1991). For crop plants, polyploidy may often lead to increased cell and organ size and potentially yield (Osborn *et al.*, 2003).

Polyploids arise constantly either by endo-reduplication of the basic set of chromosomes (Autopolyploids) or as a result of combining related but different species (Allopolyploids). Both autopolyploids and allopolyploids are common in nature, with allopolyploids predominating (Ramsey and Schemske, 1998; Soltis *et al.*, 2004a; Wendel and Doyle, 2005). From a human nutrition point of view both the forms are important (Hilu, 1993), as well as the “intermediate” types of polyploids such as segmental allopolyploids.

Polyploidization can induce certain rapid genomic and epigenomic changes and cause a restructuring of the genome. This was observed in synthetic polyploids (Comai, 2000; Wendel, 2000). Sequence elimination (Feldman *et al.*, 1997; Liu *et al.*, 1998; Ozkan *et al.*, 2001), reactivation of transposable elements (Kashkush *et al.*, 2003), changes in methylation (Shaked *et al.*, 2001) and gene expression changes (He *et al.*, 2003) were noticed upon amphiploid formation in wheat. Upon polyploidization, duplicated genes can theoretically attain three different fates: functional diversification (neo-functionalization), transcriptional gene silencing (non-functionalization) and retention of original function (sub-functionalization) (Lynch and Force, 2000). Sub-functionalization and neo-functionalization have been observed in several species (Adams *et al.*, 2003; Cusack and Wolfe, 2007; Duarte *et al.*, 2006; Force *et al.*, 1999; Liu and Adams, 2007). Both sub-functionalization and neo-functionalization are important processes from an evolutionary perspective because both processes can lead to the preservation of the members of a duplicate gene pair (Lynch and Force, 2000; Ohno, 1970). Because duplicated genes should have redundant functions immediately after they are formed, one of the copies can accumulate deleterious mutations and eventually be lost without effect on the fitness of an individual (Lynch and Conery, 2000) and sub-functionalization plays an important role shortly after gene duplication but once the age of the duplicate pair increases, neo-functionalization becomes increasingly likely and assumes a more important role later on (Ohno, 1970; Stephens, 1951). Further it has been suggested by Rastogi and Liberls, (2005) that sub-functionalization could serve as a preservational transition state leading to

neo-functionalization. Thus upon polyploidization, both sub-functionalization and neo-functionalization may play significant roles in functional diversification and duplicate gene retention.

The effect of polyploidization on transcriptional gene silencing and gene expression is an object of intense interest among researchers throughout the world (Comai, 2000; Kashkush *et al.*, 2002; Wu *et al.*, 2003; He *et al.*, 2003; Adams *et al.*, 2003; Adams *et al.*, 2004; Wang *et al.*, 2004; Hegarty *et al.*, 2005; Lai *et al.*, 2006; Tate *et al.*, 2006; Wang *et al.*, 2006b). Transcriptional gene silencing is a widespread phenomenon observed in polyploids such as cotton (Wendel, 2000), wheat (Kashkush *et al.*, 2002) and synthetic allotetraploid species of *Arabidopsis* (Comai *et al.*, 2000). Soltis *et al.*, (2004a) reported that approximately 5% of loci in *Tragopogon* and cotton were silenced as revealed by cDNA-AFLP technique and about 10% of genes were found to be repressed in *Arabidopsis* polyploids (Wang *et al.*, 2004). On a larger scale, gene expression in polyploids using microarray analysis enabled the identification of hundreds of genes that are differentially expressed between two tetraploid *Arabidopsis* lines (synthetic autotetraploid *A. thaliana* lines and *A. arenosa*, a natural allotetraploid) (Wang *et al.*, 2006b). Using SSCP analysis, Mutti JS, Gill KS (The 11th International Wheat Genetics Symposium proceedings, 2008) proposed that about 91% of the wheat genes expressed from two or three copies with only 9% showing silencing of two copies based on the sample of 652 genes indicating that most of the homoeologue genes are expressed at all times and the gene expression pattern in wheat is balanced. The relative expression levels and gene silencing of duplicated genes in polyploids can be variable in different parts of the plant also. Adams *et al.*, (2004) showed that the expression of a calmodulin-binding protein and an ubiquitin-protein ligase in the synthetic tetraploid ranged from silencing of one homoeologue in some organs to equal expression of both homoeologues in other organs. In some instances complete partitioning of expression between homoeologues can occur in different organs as was observed with the alcohol dehydrogenase A (*adhA*) gene in cotton, in which only one homoeologue is expressed in cotton petals and the other homoeologue in styles (Adams *et*

*al.*, 2004). Common/bread wheat because of its recent history is an excellent model of polypoidy and has been a good example for allopolyploidization in crop species (Kihara, 1924).

#### 1.4 Origins of A, B and D genomes of bread wheat

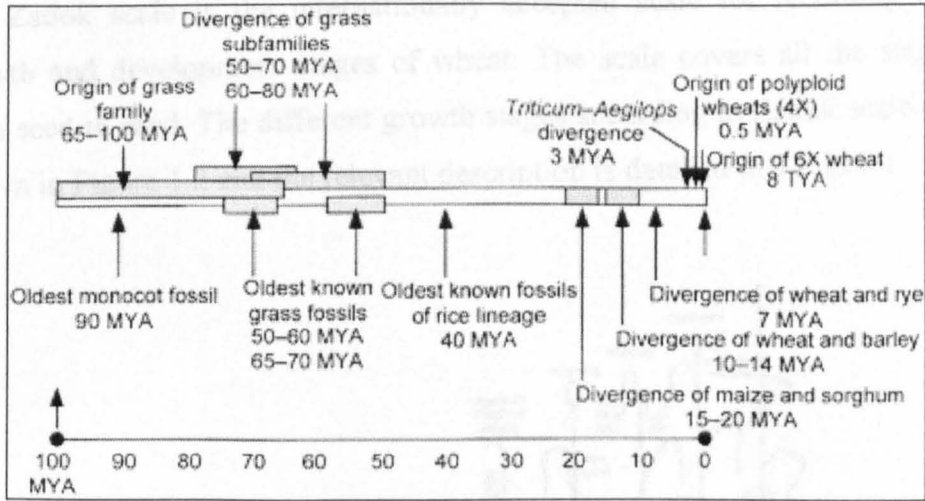
Early cytogenetic studies on polyploid wheats demonstrated that *Triticum aestivum* evolved from the hybridization of tetraploid wheat (*Triticum turgidum*) with *Aegilops tauschii* which contributed the D genome (Kihara, 1944, McFadden and Sears, 1946). Chapman *et al.*, (1976) determined that the A genome in *Triticum turgidum* has been contributed by *Triticum uratu*. The second genome in *Triticum turgidum* is the B genome which may have been contributed by *Triticum speltoides* (Dvork and Zhang, 1990). Later on Cox (1998) summarized the believed evolution of bread wheat which is now widely accepted. According to Cox (1998), Common wheat or bread wheat is an allohexaploid (genomic formula AABBDD according to Waines and Barnhart, 1992) derived through hybridization between a domesticated form of tetraploid wheat, wild emmer (*Triticum turgidum* ssp. *dicoccoides* genomic constitution AABB) and the diploid *Aegilops tauschii* (genomic constitution DD). Wild emmer itself is supposed to be an allotetraploid derived through hybridization between two wild diploids; *Triticum uratu* contributing the A genome and probably an *Aegilops speltoides* relative which is no longer extant, contributing the B genome. However, despite decades of intensive research, the origin of the B genome has remained inconclusive (Huang *et al.*, 2002a). Recently Peterson *et al.*, (2006) based on a phylogenetic analysis of the single-copy nuclear genes *DMC1* and *EF-G* from the three genomes in the hexaploid wheat species were able to provide for the first time strong molecular supporting evidence of the D genome being derived from *Aegilops tauschii*, the A genome being derived from *Triticum uratu*, and the hitherto enigmatic B genome being derived from *Aegilops speltoides*.

Huang *et al.*, (2002) by analyzing the *Acc-1* (plastid acetyl-CoA carboxylase) and *Pgk-1* (plastid 3-phosphoglycerate kinase) genes

established the timeline of wheat evolution based on gene sequence comparisons of *Triticum* and *Aegilops* species. According to Huang *et al.*, (2002) the A genome of polyploid wheat diverged from *T. urartu* less than half a million years ago (MYA), indicating a relatively recent origin of polyploid wheat. The D genome sequences of *T. aestivum* and *Ae. tauschii* are similar, confirming that *T. aestivum* arose from hybridization of *T. turgidum* and *Ae. tauschii* around 8,000 years ago (Figure 1.1).

Common wheat or bread wheat is a segmental allopolyploid (6x) composed of three genomes of seven groups of chromosomes. On the basis of genetic similarities, the 21 pairs of chromosomes fall into seven homoeologous groups, each containing one pair of chromosomes from the A, B, and D genomes (Sears, 1954). The term homoeologous refers to a pair of chromosomes that have alleles for the same gene. Because of this homoeology in hexaploid wheat (AABBDD), a range of chromosomal abnormalities (aneuploidy) are able to survive, which would not be able to survive in a diploid species such as barley (*Hordeum vulgare* L.) and maize (*Zea mays* L.). Since the loss of a pair of chromosomes can be compensated by two additional doses of a homoeologue, 42 compensating nulli-tetrasomics lines were developed by Sears (1966). Sears (1954) also described the effects of aneuploidy of each chromosome including nullisomic and monosomic lines. The monosomic chromosomes fail to divide properly and this property was exploited to produce a series of chromosome-arm aneuploids such as monotelosomics, ditelosomics, tritelosomics, and iso-chromosome lines (Sears and Sears, 1978). Among cultivated crops, hexaploid wheat has the largest genome size at 17 GB, ~8-fold larger than that of Maize and 40-fold larger than that of Rice (Arumuganthan and Earle, 1991).

Figure 1.1 Recent co-evolutionary history of cereals

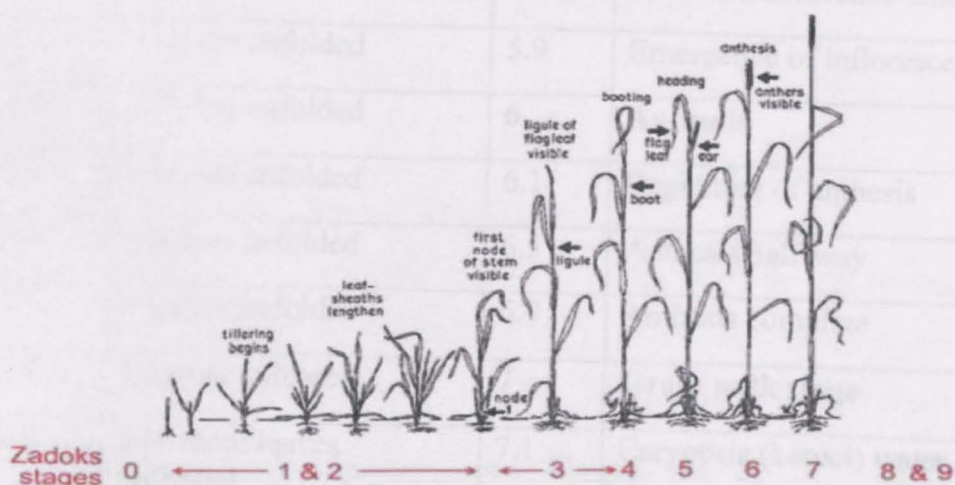


**Figure 1.1** Recent co-evolutionary history of cereals. In the figure the origin of wheat can be seen a recent event. (Figure courtesy of W. J. Raupp, based on discussions with P. F. Byrne, Colorado State University, Fort Collins, with additional data from Huang *et al.*, (2002).



### 1.5 Different growth stages of wheat

The Zadok scale is the internationally accepted scale for recording the growth and development stages of wheat. The scale covers all the stages from seed to seed. The different growth stages according to Zadok scale are shown in Figure 1.2 and the relevant description is detailed in Table 1.1



**Figure 1.2** Schematic diagram of the different growth and developmental Zadok stages of the wheat plant. (Source: Zadok, J.C., Chang, T.T., Konzak, 1974).

**Table 1.1** Different growth stages of wheat along with the description of each stage.

Zadok Stage	Description	Zadok Stage	Description
0	Seed germination	4	Booting
0.0	Dry seed	4.1	Flag leaf sheath extending
0.1	Start of imbibition (water absorption)	4.3	Boots just visibly swollen
0.3	Imbibition complete	4.5	Boots swollen
0.5	Radicle (root) emerged from caryopsis (seed)	4.7	Flag leaf sheath opening
0.7	Coleoptile	4.9	First awns visible

0.9	Leaf just at coleoptile tip	5	<b>Heading</b>
<b>1</b>	<b>Main stem leaf production</b>	5.1	First spikelet of inflorescence just visible
1.0	First leaf through coleoptile	5.3	1/4 of inflorescence emerged
1.1	First leaf unfolded	5.5	1/2 of inflorescence emerged
1.2	2 leaves unfolded	5.7	3/4 of inflorescence emerged
1.3	3 leaves unfolded	5.9	Emergence of inflorescence
1.4	4 leaves unfolded	6	<b>Anthesis</b>
1.5	5 leaves unfolded	6.1	Beginning of anthesis
1.6	6 leaves unfolded	6.5	Anthesis half-way
1.7	7 leaves unfolded	6.9	Anthesis complete
1.8	8 leaves unfolded	7	<b>Grain milk stage</b>
1.9	9 or more leaves unfolded	7.1	Caryopsis (kernel) water ripe
<b>2</b>	<b>Tiller production</b>	7.3	Early milk
2.0	Main shoot only	7.5	Medium milk
2.1	Main shoot and 1 tiller	7.7	Late milk
2.2	Main shoot and 2 tillers	8	<b>Grain dough stage</b>
2.3	Main shoot and 3 tillers	8.3	Early dough
2.4	Main shoot and 4 tillers	8.5	Soft dough
2.5	Main shoot and 5 tillers	8.7	Hard dough
2.6	Main shoot and 6 tillers	9	<b>Ripening</b>
2.7	Main shoot and 7 tillers	9.1	Caryopsis hard (difficult to divide)
2.8	Main shoot and 8 tillers	9.2	Caryopsis hard (not dented by thumbnail)
2.9	Main shoot and 9 or more tillers	9.3	Caryopsis loosening in daytime
<b>3</b>	<b>Stem elongation</b>	9.4	Over-ripe, straw dead and collapsing
3.0	Pseudostem (leaf sheath) erection	9.5	Seed dormant
3.1	First node detectable	9.6	Viable seed giving 50% germination
3.2	2nd node detectable	9.7	Seed not dormant

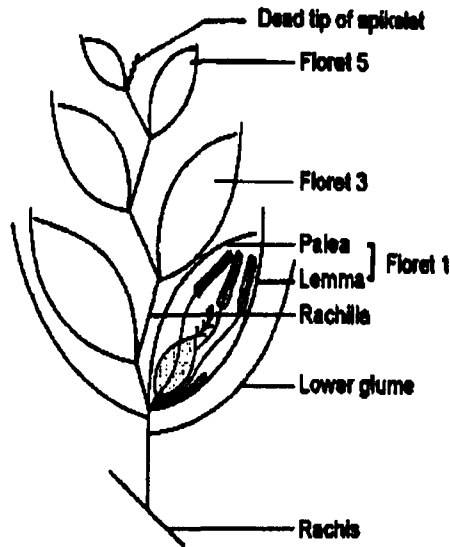
3.3	3rd node detectable	9.8	Secondary dormancy induced
3.4	4th node detectable	9.9	Secondary dormancy lost
3.5	5th node detectable		
3.6	6th node detectable		
3.7	Flag leaf just visible		
3.9	Flag leaf ligule just visible		

### 1.6 Reproductive Biology of wheat

Wheat is predominantly a self pollinated crop, however outcrossing can occur at 10% or above depending upon the population density, genotype and environmental conditions (Jain, 1975). The wheat inflorescence is a determinate, composite spike composed of main axis (rachis) bearing spikelets separated by short internodes. The majority of flowers on the wheat spike are hermaphroditic, although some are unisexual (De Vries, 1971). Each spikelet comprises an axis (rachilla), which bears two glumes and each ear consists of twenty or more spikelets that have a defined sequence of maturity (Schwarzcher, 1997). The floret contains two sheathing structures called the outer lemma and the inner palea; these envelope two lodicules, three anthers and the ovary (Figure 1.3). Within each of the two major florets of each spikelet (one slightly younger than the other) there are three synchronously developing anthers (Bennett *et al.*, 1971).

Anthesis occurs for about three to ten days depending on the environment after the ear emerges from the flag leaf sheath. Anthesis occurs first in floret 1 of the spikelet of the upper two-thirds of the ear. It progresses to the first floret of the basal spikelet and to the second floret of the upper spikelet on the next day. This progression continues so that the third and fourth florets in the basal spikelet are the last in which anthesis occurs (Evans *et al.*, 1972). Often, the higher order florets (three to five), although going through the process of anthesis and becoming pollinated, do not produce grain. The

basal florets are generally fertile, but some of the distal florets die sequentially during ear development.



**Figure 1.3** Diagram of wheat spikelet displaying the arrangement of the florets inside (Source: Kirby and Appleyard, 1987)

### 1.7 General information of *Arabidopsis thaliana*

*Arabidopsis thaliana* is a small dicotyledonous flowering plant belonging to the family of Brassicaceae to which many economically important crops such as turnip, raddish, cabbage, broccoli and canola belong.

#### 1.7.1 *Arabidopsis* as a model plant

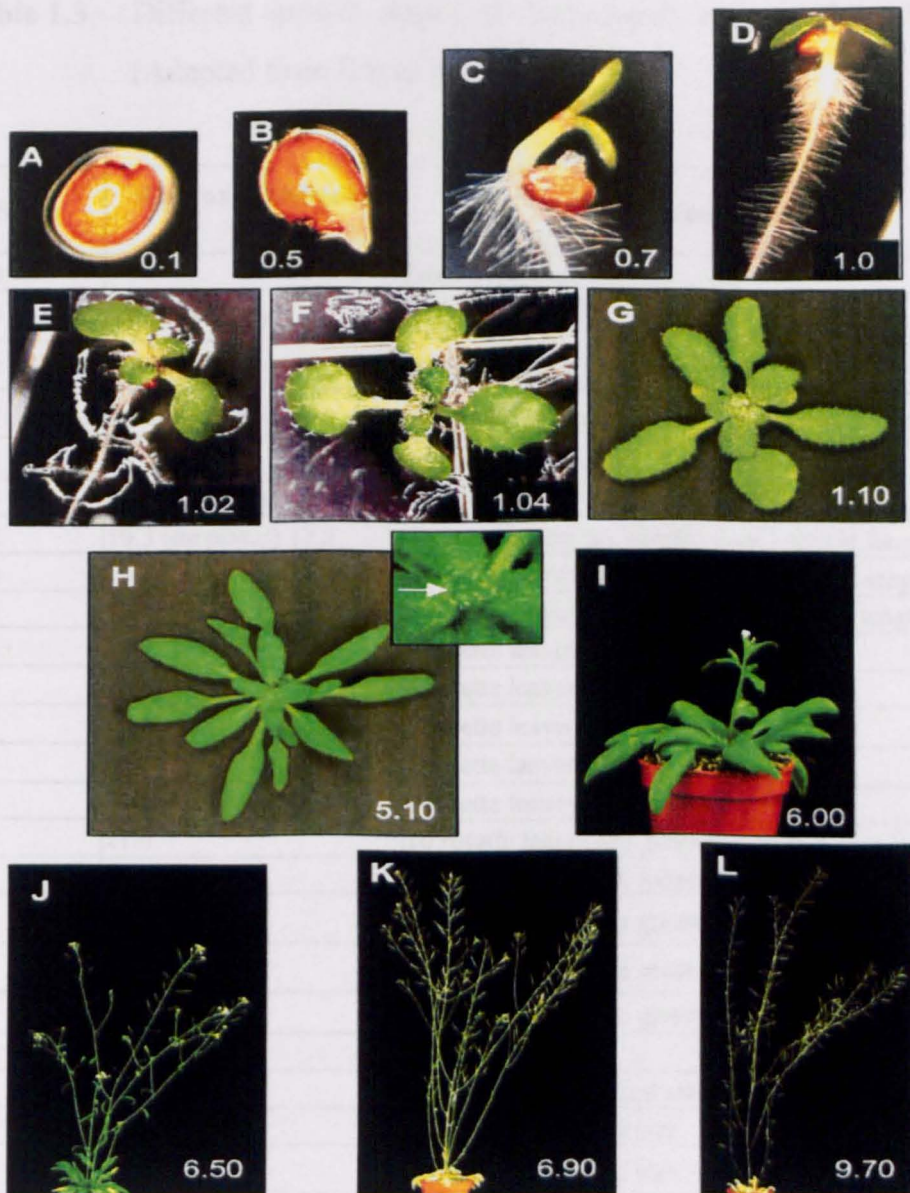
Although *Arabidopsis* is not an economically important crop it has been focus of intense genetic, biochemical and physiological study for more than 40 years because of several traits which makes it useful for laboratory studies and make it a model plant (Rensink *et al.*, 2005). *Arabidopsis* is a very small plant and has very limited space requirements, it has a fast life cycle (6-8 weeks from seed to seed) which means that it produces very large numbers of progeny in a short span of time. It is also almost entirely self-fertilizing and it can easily be grown in greenhouses or in indoor growth chambers. From the molecular biologist's point of view it has a relatively

small, diploid, genetically tractable genome that has been sequenced completely (Arabidopsis Genome Initiative, 2000). It can be easily manipulated by genetic engineering compared to many other plant genomes and has a large collection of mutant lines generated by X-ray irradiation, chemical mutagenesis and insertional mutagenesis using T-DNA and transposons for reverse genetics.

### **1.7.2 Growth stages of Arabidopsis**

The different growth and developmental stages (Figure 1.4) and the timetable of the growth stages (Table 1.2) of *Arabidopsis* ecotype Columbia-O have been determined by Boyes *et al.*, (2001). The growth stages were developed according to BBCH scale (Lancashire *et al.*, 1991) and all together there are 30 growth stages in *Arabidopsis* which cover the development of the plant from seed imbibition to the completion of flowering and seed maturation. The stages and timeline are for plants grown under the following conditions: Columbia ecotype plants were grown in soil at in 16-hour days/ 8-hour nights with temperatures of 22°C during the day and 20°C at night.





**Figure 1.4** Different growth and development stages of *Arabidopsis*. (A) Stage 0.1, imbibition. (B) Stage 0.5, radicle emergence. (C) Stage 0.7, hypocotyl and cotyledons emerged from seed coat. (D) Stage 1.0, cotyledons opened fully. (E) Stage 1.02, two rosette leaves\_ 1 mm in length. (F) Stage 1.04, four rosette leaves\_ 1 mm in length. (G) Stage 1.10, ten rosette leaves\_ 1 mm in length. (H) Stage 5.10, first flower buds visible (indicated by arrow in inset). (I) Stage 6.00, first flower open. (J) Stage 6.50, mid flowering. (K) Stage 6.90, flowering complete. (L) Stage 9.70, senescent and ready for seed harvest.

**Table 1.3** Different growth stages of *Arabidopsis* ecotype Columbia-O  
(Adapted from Boyes *et al.*, 2001)

Stage	Approx. Number of days	Description of stage
<b>0.00</b>		<b>Seed germination</b>
0.10	3.0 (on plates)	Seed imbibitions
0.50	4.3 (on plates)	Radicle emerges from seed coat
0.70	5.5 (on plates)	Hypocotyl and cotyledon emerge from seed coat
<b>1.00</b>		<b>Rosette growth</b>
1.00	6.0 (on plates)	Cotyledons fully open
1.02	10.3 (on plates) 12.5	2 rosette leaves are greater than 1 mm in length
1.03	14.4 (on plates) 15.9	3 rosette leaves are greater than 1 mm in length
1.04	16.5	4 rosette leaves are greater than 1mm in length
1.05	17.7	5 rosette leaves are greater than 1mm
1.06	18.4	6 rosette leaves are greater than 1mm
1.07	19.4	7 rosette leaves are greater than 1mm
1.08	20	8 rosette leaves are greater than 1mm
1.09	21.1	9 rosette leaves are greater than 1mm
1.10	21.6	10 rosette leaves are greater than 1mm
1.11	22.2	11 rosette leaves are greater than 1mm
1.12	23.3	12 rosette leaves are greater than 1mm
1.13	24.8	13 rosette leaves are greater than 1mm
1.14	25.5	14 rosette leaves are greater than 1mm
<b>3.00</b>		<b>Rosette Growth</b>
3.20	18.9	Rosette is 20% of final size
3.50	24	Rosette is 50% final size
3.70	27.4	Rosette is 70% final size
3.90	29.3	Rosette growth is complete
<b>5.00</b>		<b>Inflorescence emergence</b>
5.10	26	First flower buds are visible in the rosette, plant has not yet bolted
<b>6.00</b>		<b>Flower production</b>
6.00	31.8	First flower is open, petals are at 90 degree angle to the pistil
6.10	35.9	10% flowers to be produced are open
6.30	40.1	30% flowers to be produced are open
6.50	43.5	50 % flowers to be produced are open
6.90	49.4	Flowering complete, flowers are no longer produced.
8.00		Silique or fruit ripening. Seed pods become brown and then shatter.
<b>8.00</b>	<b>48</b>	<b>First silique or seed pod shatters.</b>
9.00		Whole plant senescence begins. Plant starts to lose, pigment becoming brownish.
<b>9.70</b>		<b>Senescence complete</b>

## 1.8 Meiosis

### 1.8.1 An overview

The term 'meiosis' (from the Greek word *meioun*, means to diminish) was first introduced by Farmer and Moore (1905) to describe a specialized form of cell division. Meiosis is central to all sexually reproducing forms of living organisms and is highly conserved among them. During meiosis a single round of DNA replication is followed by two rounds of cell division (Meiosis I and Meiosis II) resulting in formation of four daughter cells each having half the chromosome number of the parent. The first nuclear division results in the segregation of homologous chromosomes to opposite poles and is unique to meiosis. During the second nuclear division, sister chromatids are separated resulting in the formation of four haploid nuclei and following meiotic cytokinesis results in the formation of haploid cells.

In contrast, during mitosis (after DNA synthesis; S phase) sister chromatids which are joined at the centromeres line up to the spindle fibers with their kinetochores pointing towards opposite poles. The contraction of spindle fibers proceeds to separate the sister chromatids to opposite poles. Formation of the nuclear membrane around each chromosome results in the formation of individual cells each having a copy of the maternal and a copy of the paternal chromosomes. The final result of mitosis is the formation of two identical cells from one cell.

Meiosis produces four gametes each having half of the chromosome complement of the parent. During fertilization and subsequent events, the gametes from both the parents fuse and thereby produce a diploid complement of chromosomes within the cell. A considerable amount of information is available on the events occurring during meiosis in a number of eukaryotes including yeast (Roeder, 1995), *Drosophila melanogaster* (Orr-Weaver, 1995) and plants (Dawe, 1998).



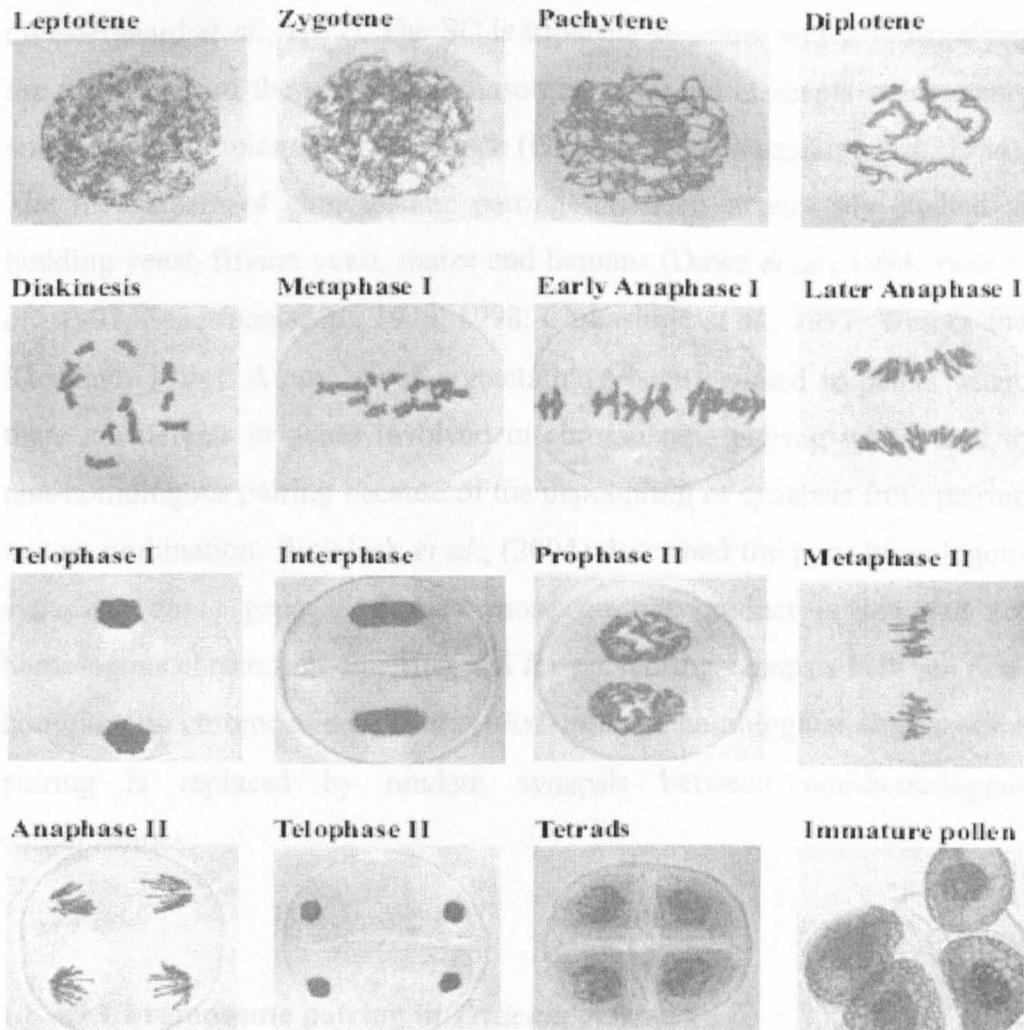
### 1.8.2 The anthers and stages of meiosis in wheat

In wheat, the anther has four lodicules containing archesporial tissue, and it is the cells of each archesporium that differentiate into the central core of meiocytes (Dickinsons, 1981). Around 200 to 300 meiocytes are surrounded by a layer of cells called the tapetum which nourish and secure the anther (Bennett *et al.*, 1971). Among all higher plants, wheat has got one of the shortest meiotic division cycles. Meiosis continues for around 24 hours at 20°C (Bennett *et al.*, 1973). Bennett (1971) found a 42 hour meiotic cycle in the diploid *Triticum monococcum* and a 30 hour meiotic cycle in the tetraploid *Triticum dicoccum* indicating that as the ploidy level increases, the duration of meiosis decreases. The stage of meiosis in any particular anther can be determined approximately from its size, the size of the florets, and its position with respect to other florets in a spikelet in wheat.

Meiosis consists of two meiotic stages- Meiosis I and II, with each comprising of four stages; prophase I, metaphase I, anaphase I and telophase I, followed by prophase II, metaphase II, anaphase II and telophase II (Figure 1.5). Prophase I is the most complex and the slowest. It is further sub-divided into five stages: leptotene, zygotene, pachytene, diplotene and diakinesis. Leptotene is the longest phase during prophase I, lasting approximately 10 hours in wheat (Bennett *et al.*, 1973) and it is at this time that the chromosomes become visible. Leptotene is followed by zygotene and continues for approximately three to four hours (Bennett *et al.*, 1973). The chromosome telomeres orient in a unique arrangement called a “bouquet”. During zygotene, homologous chromosomes synapse along their length. Homoeologous multivalents are frequently observed during pairing in zygotene (Sears, 1977; Roberts *et al.*, 1999). The chromosomes start to condense during pachytene and the chromosome complement is represented by a haploid number of bivalents, during normal meiosis. During this stage the genetically important phenomenon of crossing-over between non-sister chromatids occurs. The pachytene stage lasts for approximately two hours in wheat (Bennett *et al.*, 1973). The synapsed chromosomes begin to separate during diplotene and two of the four chromatids from each bivalent remain attached via chiasmata which represent the points of physical crossing over.

During diakinesis, chromosomes reach their maximum condensation. Towards the end, spindle formation is initiated and the nuclear envelope breaks down. Both diplotene and diakinesis collectively last for one hour in wheat (Bennett *et al.*, 1973).

Prophase I is followed by two successive nuclear divisions resulting in the production of four haploid cells with the haploid complement of chromosomes that have been subjected to genetic recombination.



**Figure 1.5** Different stages of meiosis in *Lilium regale* (Modified from McLeish and Snoad, 1958).

### 1.8.3 Chromosome pairing during meiosis in eukaryotes

Chromosome pairing is one of the important processes occurring during Meiosis I and is essential for efficient genetic recombination and proper segregation of half of the bivalents during Anaphase I. The homologous chromosomes (termed homologues) start in a premeiotic somatic nucleus randomly organized with respect to each other but end up during the pachytene stage of meiotic prophase in close association. The process involves reciprocal recognition, co-alignment and synapsis. The intimate association of the homologues, which are composed of sister chromatids is facilitated by a proteins structure called the Synaptonemal Complex (SC)

(Westergaard *et al.*, 1972). The SC is tripartite structure which forms along the full length of the paired chromosomes as lateral elements at leptotene, and reaches completion by pachytene (Gillies, 1984; Wettstein *et al.*, 1984). The mechanism of chromosome pairing has been extensively studied in budding yeast, fission yeast, maize and humans (Dawe *et al.*, 1994; Bass *et al.*, 1997; Scherthan *et al.*, 1996, 1998; Chikashige *et al.*, 1997; Weiner and Kleckner, 1994). A number of mutants have been isolated in plants where there are defects in genes involved in chromosome pairing which lead to non-homologous pairing because of the uncoupling of synapsis from pairing and recombination. Wojciech *et al.*, (2004) described the poor homologous synapsis (*phs1*) gene in maize whose protein product is required for homologous chromosome pairing and for preventing synapsis between non-homologous chromosomes. In the *phs1* mutant, homologous chromosome pairing is replaced by random synapsis between non-homologous chromosomes.

#### **1.8.3.1 Chromosome pairing in *Triticum aestivum***

Even though wheat is a polyploid species pairing and recombination are confined to strictly homologous chromosomes during meiosis I despite the coexistence in the genome of three homoeologous chromosomes. The genetic control of chromosome pairing in wheat is dependant on a series of suppressing and promoting *Pairing homoeologous (Ph1)* genes (Sears, 1976). The major locus controlling this behavior in hexaploid wheat is the *Ph1* gene (Riley and Chapman, 1958; Sears, 1976). The *Ph1* was originally identified by its ability to affects chromosome pairing in wheat mutants. The original *Ph1* locus was defined by a single deletion of 70 Mb containing over 200 genes (Gill *et al.*, 1993, 1996) but it has been delimited to a region containing less than 7 genes (Sears, 1977; Gill *et al.*, 1993; Roberts *et al.*, 1999). The *Ph1* gene was mapped and found to be located on the long arm of chromosome 5B linkage group (Riley and Chapman, 1958; Sears and Okamoto, 1958). *Ph1* acts as a dominant gene suppressing pairing between homoeologous chromosomes while allowing pairing between homologous chromosomes (Riley and Chapman, 1958). In the presence of *Ph1* gene, true

homologues are synchronized in their chromatin conformational changes and the individual genomes are localized within the cell, before pairing with each other, preventing pairing to other homoeologous chromosomes. But in the absence of *Ph1*, any chromosome can remodel without the requirement for the existence of an identical or near identical chromosome and thereby increasing the chance of it pairing with a homoeologue or a related chromosome rather than pairing between true homologues.

Mutants carrying deletions in the *Ph1* locus (*Ph1b*) exhibit a degree of multivalent formation at metaphase I due to the pairing of homoeologous chromosomes (Sears, 1977; Roberts *et al.*, 1999). But considerable numbers of bivalents are still observed during metaphase I in the *Ph1b* mutant and the mutant is stable for many generations. Polyploidization may have played a role in the origin of *Ph1* locus because the *Ph1* locus is absent in diploid ancestors of wheat (Chapman and Riley, 1970). Because of the complexity of the *Ph1* locus a two-part strategy was devised to dissect its molecular nature. The first part is based using the conservation of gene order between the highly repetitive, 17,000 Mb hexaploid wheat genome and the smaller rice and *Brachypodium sylvaticum* genomes, which provided markers to saturate the *Ph1* locus in wheat. The second part of the strategy was based on availability of new deletion lines to dissect the *Ph1* locus physically. Using these approaches, the *Ph1* locus was confined to a 2.5 Mb region of wheat chromosome 5B containing a structure consisting of a segment of sub-telomeric heterochromatin that inserted into a cluster of *cdc2* (cdk)-related genes following polyploidization (Griffiths *et al.*, 2006).

In addition to *Ph1* two further suppressors of homoeologous chromosome pairing have been identified in wheat (for reviews, see Sears 1976). The locus located on short arm of chromosome 3D (3DS), termed *Ph2*, controls chromosome pairing in wheat hybrids with alien species (Mello-Sampayo, 1971; Mello-Sampayo and Lorente, 1968; Upadhy and Swanminathan, 1967). A deletion mutant of *Ph2* locus, *ph2a* (Sears 1982) and an EMS-induced mutant *ph2b* (Wall *et al.*, 1971) showed increased levels of homoeologous recombination in hybrids with *Triticum kotshyi* var.

*variabilis* or rye (*Secale cereale*). Recently Dong *et al.*, (2002) isolated *TaMSH7* gene, a wheat homologue of yeast DNA mismatch repair gene, which is located in the region deleted in ph2a region in wheat which showed altered recombination frequency in inter-specific hybrids. The second locus is a gene suppressor of smaller magnitude located on short arm of chromosome 3A (3AS) (Driscoll, 1972; Mello-Sampayo and Canas, 1973).

Despite the importance of the *Ph1* genes in controlling homologous chromosome pairing and synapsis, there are no reports of the role of any proteins that have been shown to have a role in homologous chromosome pairing and synapsis. Recently *TaASY1* gene, a homologue of *ScHOP1*, *AtASY1* and *OsPAIR2* isolated from bread wheat were shown to be involved in pairing and synapsis of homologous chromosomes (Boden *et al.*, 2007). It was shown that the absence of *Ph1* causes elevated levels of transcripts of *TaASY1* to accumulate and this suggests that it plays a role in homologous chromosome pairing (Boden *et al.*, 2008).

## **1.9 Homologous Recombination (HR)**

### **1.9.1 An overview**

HR is the process of exchange (“crossing over”) or replacement (“gene conversion”) of a DNA region by its homologous DNA sequence from the homologous chromosome or the sister chromatid. HR is an essential process involved in genome integrity through accurate repair of DNA damage and it also generates novel combinations of genetic diversity through crossing over during meiosis.

### **1.9.2 HR at meiosis**

While HR is required for maintaining the integrity of genome in somatic cells, HR at meiosis (termed meiotic recombination; meiHR) has evolved to create genetic variation through crossing over and it also plays a mechanical role ensuring physical connections in the form of chiasmata that allow the chromosomes to orient at right angles to the spindle fibers and thus

segregate accurately during the first division of meiosis. Because meiosis is a specialized form of cell division and the meiotic recombination machinery appears mostly derived from the machinery responsible for mitotic recombination and DNA repair in somatic cells, both the cell divisions share many components. HR has been extensively studied in yeast due to the ability to recover and analyze all the products of single meiosis event as tetrads.

### **1.9.3 Models of HR**

A number of models have been proposed to explain how HR actually occurs in a cell. They include the Holliday model (Holliday, 1964), Meselson-Radding model (Meselson-Radding, 1975) and the double strand break repair model (Szostak *et al.*, 1983). These are explained briefly below.

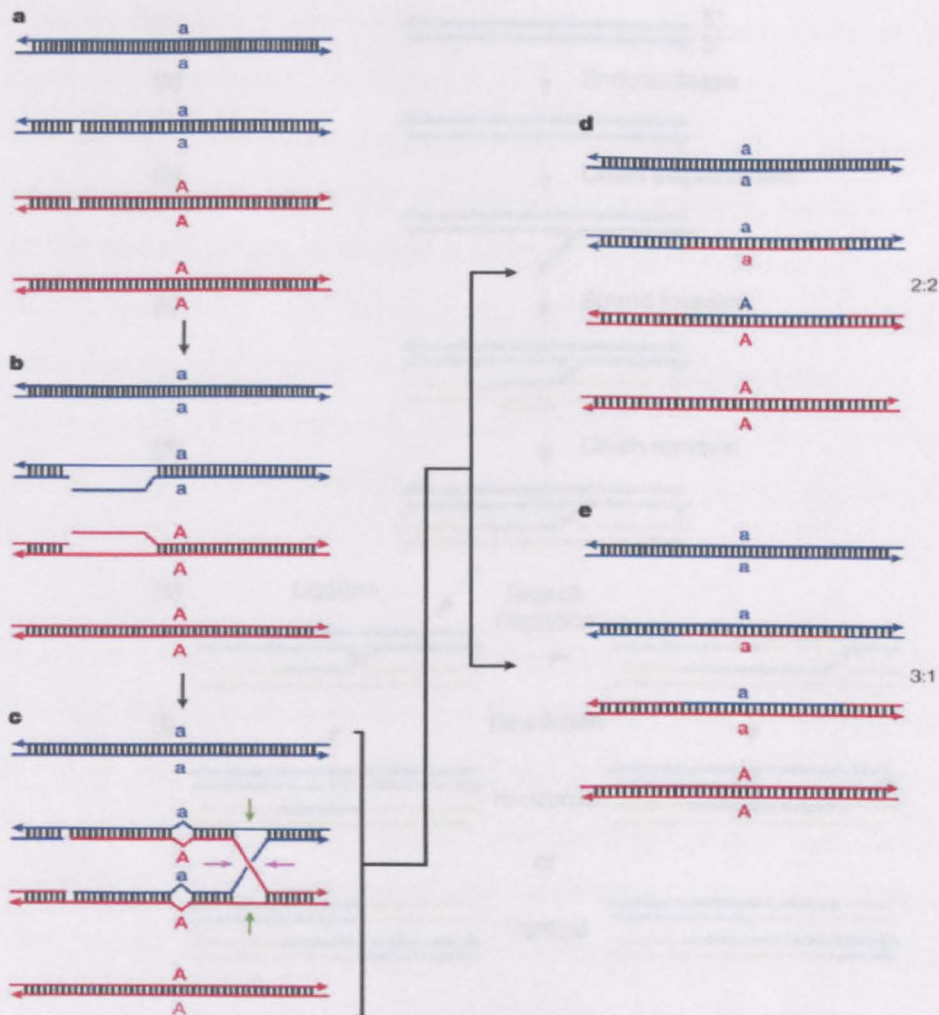
#### **1.9.3.1 The Holliday and Meselson-Radding models**

For a number of years, the genetic recombination model put forth by Holliday (1964), Holliday and Halliwell (1968) has been a popular model for explaining the relationships between aberrant segregation and crossing over. Following formation of single-stranded nicks in each of the DNA strands, exchange of DNA strands generates a region of symmetrical heteroduplex DNA on each of the two interacting sister strands, giving rise to a crossed strand structure known as a Holliday junction (Figure 1.6). The resultant Holliday junction is resolved in either one of two specific orientations (Figure 1.6a and 1.6b) to yield recombinant progeny and the outcome of these steps is the generation of symmetric recombinant DNA molecules that are either "spliced" or "patched."

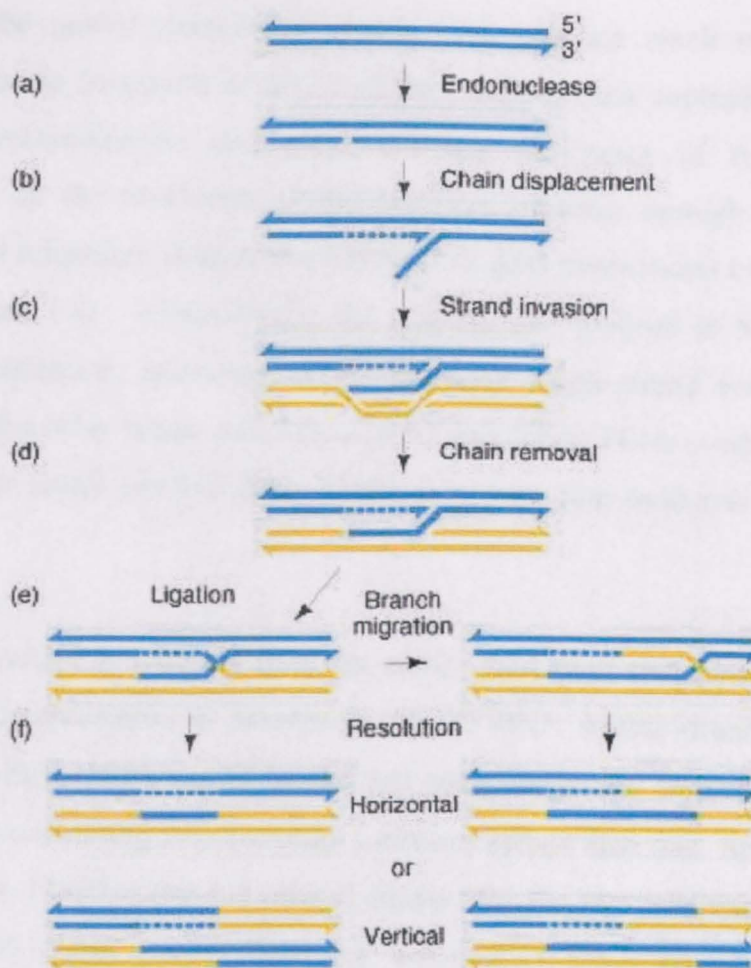
The Holliday model was able to explain several aspects of recombination including several types of aberrant segregation and the association of segregation ratios with cross-overs. One of the assumptions of Holliday model was that heterduplex DNA occurred equally between both chromatids. But studies conducted in *Ascobolus* by Stadler and Towe (1971)

found that heteroduplex DNA occurs only on one of the chromatids in this species. To account for this observation Meselson-Radding (1975), proposed a recombination model that is based on the formation of an asymmetric heteroduplex DNA. In this model, recombination is initiated by the transfer of a single-strand segment from one strand to another, in the process displacing the homologue and leading to the formation of asymmetric heteroduplex DNA. The 3' end of the nicked strand acts as a primer from the complementary strand and DNA synthesis completely displaces the strand ahead of it (Figure 1.7). In this way the recipient strand will have heteroduplex DNA while the donor duplex will remain as homoduplex. Resolution of the heteroduplex DNA structure resembles the Holliday model. The Meselson-Radding model was able to explain the discrepancies observed in *Asobolous* recombination. Due to certain constraints (such as the parity of gene conversion for all classes of mutations in yeast and the observation of the fact that the strand upon which initiation occurs appears to be the recipient of genetic information) Stahl (1979) proposed a simpler recombination model – “the Double Strand Break Repair model”.





**Figure 1.6** The Holliday model of homologous recombination. The different steps involved are recombination initiation (a), Strand exchange (b), symmetric resolution in two possible orientations (indicated by purple and green arrows) (c) and the production of gene conversion events (d & e) (source: Liu and West, 2004)



**Figure 1.7** Meselson-Radding model of recombination (Meselson-Radding, 1975).

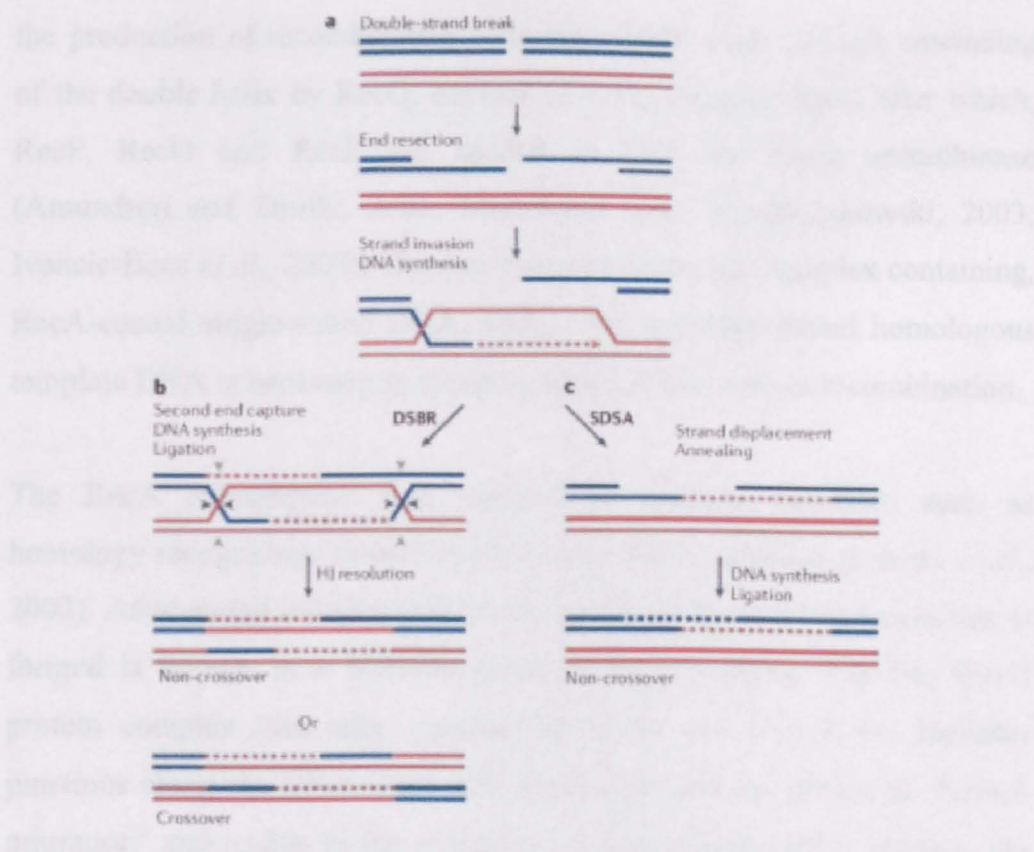
### 1.9.3.2 The Double-Strand Break Repair (DSBR) and Synthesis Dependant Strand Annealing model (SDSA) models

Two primary models that explain how HR aids DSB repair in DNA are DSBR pathway (Szostak *et al.*, 1983) and the SDSA pathway (Sung *et al.*, 2006). In both the models, recombination is initiated by a double-strand break introduced by an endonuclease into the recipient chromatid. The cut is then enlarged into a gap and with the help of other endonucleases, 3' single-stranded termini are produced. One of the free 3' ends then invades the homologous region of the donor duplex, displaces one of the strands and base pairs with the complementary single strand sequence. As DNA synthesis proceeds, with the invading strand as the primer, branch migration

displaces the newly synthesized strands. The ultimate result of branch migration is the formation of two Holliday junctions, one representing the point of recombination and other defining the point of resolution. Depending on the resolution of the Holliday junction through separate cuttings and religation, reciprocal crossovers or gene conversions events can occur (Figure 1.8). Alternatively, the reaction can proceed to SDSA by strand displacement, annealing of the extended single-strand end to the ssDNA on the other break end, followed by gap-filling DNA synthesis and ligation. The repair product from SDSA is always non-crossover (Figure 1.8).

The DSBR model is different from the earlier models of recombination in two ways: recombination is initiated by double DNA double-strand breaks and not as single DNA strand breaks and such breaks are repaired by an intermediate containing two Holliday junctions rather than one. Schwacha and Kleckner, (1995) provided critical support for the two new features of double strand repair model from the analysis of DNA recombination intermediates at a meiotic recombination hot spot using 2D gel electrophoresis.





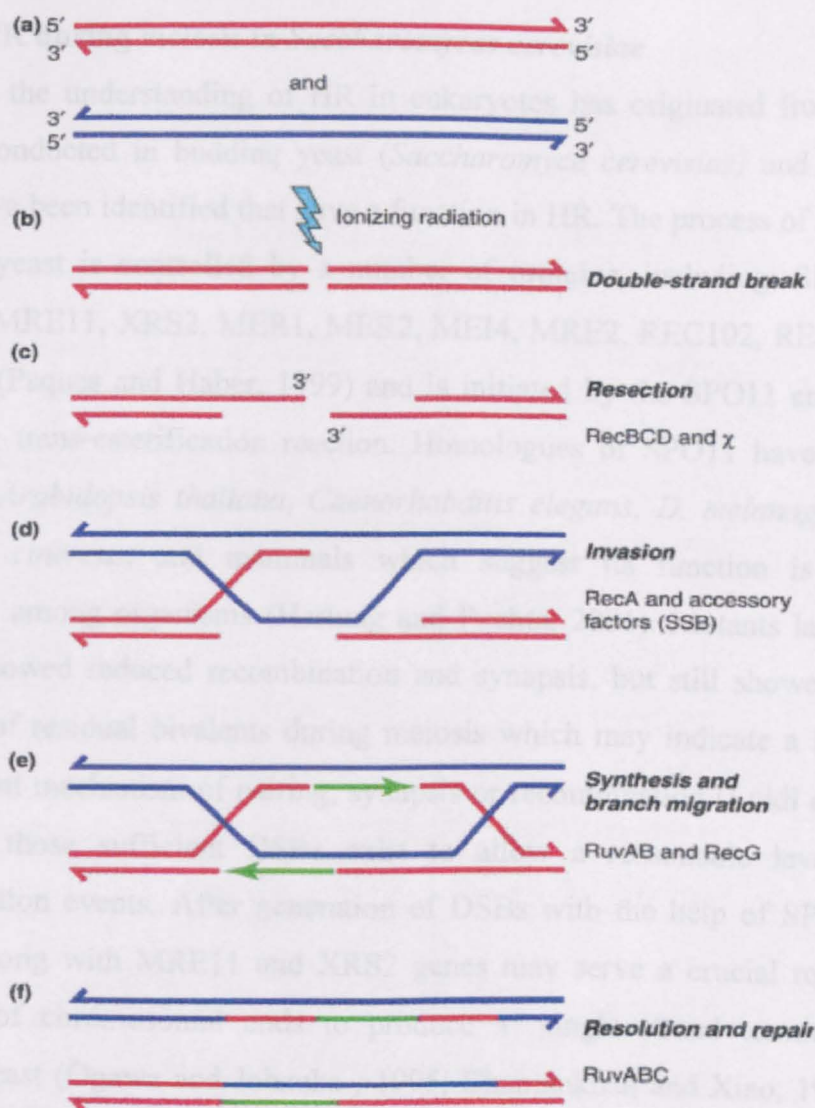
**Figure 1.8** Double Strand Break Repair and Synthesis Dependant Strand Annealing model of Homologous Recombination (Sung *et al.*, 2006)

#### 1.9.4 HR in *E. coli*

HR was first studied in *E. coli* and over the last 20 years, *E. coli* has become a good model for the study of HR. In *E. coli* normal levels of recombination requires almost 20 gene products, including; *recA*, *recB*, *recC*, *recD*, *recF*, *recG*, *recJ*, *recN*, *recO*, *recQ*, *recR*, *ruvA*, *ruvB*, *ruvC* and *ssb* genes. Many of these genes are involved at various stages of recombination and play well defined roles such as: initiation; DNA-DNA interactions; branch migration and Holliday junction resolution (Figure 1.9). In *E. coli* HR is initiated by a DSB and the breaks are processed to expose single-stranded 3' overhangs onto which RecA gene can be loaded in two different ways. In the first method, a RecBCD a heterotrimeric protein complex with DNA helicase, ssDNA and dsDNA nuclease activities coordinates loading of the RecA recombinase. In the second method a sequence of events leading to

the production of recombination competent DNA ends through unwinding of the double helix by RecQ, nucleolytic processing by RecJ, after which, RecF, RecO and RecR are needed to load the RecA recombinase (Amundsen and Smith., 2003; Morimatsu and Kowalczykowski, 2003; Ivancic-Bace *et al.*, 2003). This pre-synapsis molecular complex containing, RecA-coated single-strand DNA, along with a double-strand homologous template DNA is necessary in synapsis which is next step in recombination.

The RecA recombinase was reported to mediate processes such as homology recognition, strand invasion and strand exchange (Lusetti *et al.*, 2002). After strand invasion and strand exchange the joint molecule that is formed is known as a Holliday junction. RuvAB along with the RuvG protein complex then acts a molecular motor and moves the Holliday junctions along the DNA in an ATP-dependent process known as "branch migration" and results in the extension of heteroduplex DNA between the recombining DNA molecules (West, 2003). The RuvC protein is a resolvase protein that interacts with the RuvAB protein complex and nicks strands at the junction, thereby releasing the recombining partners from each others.



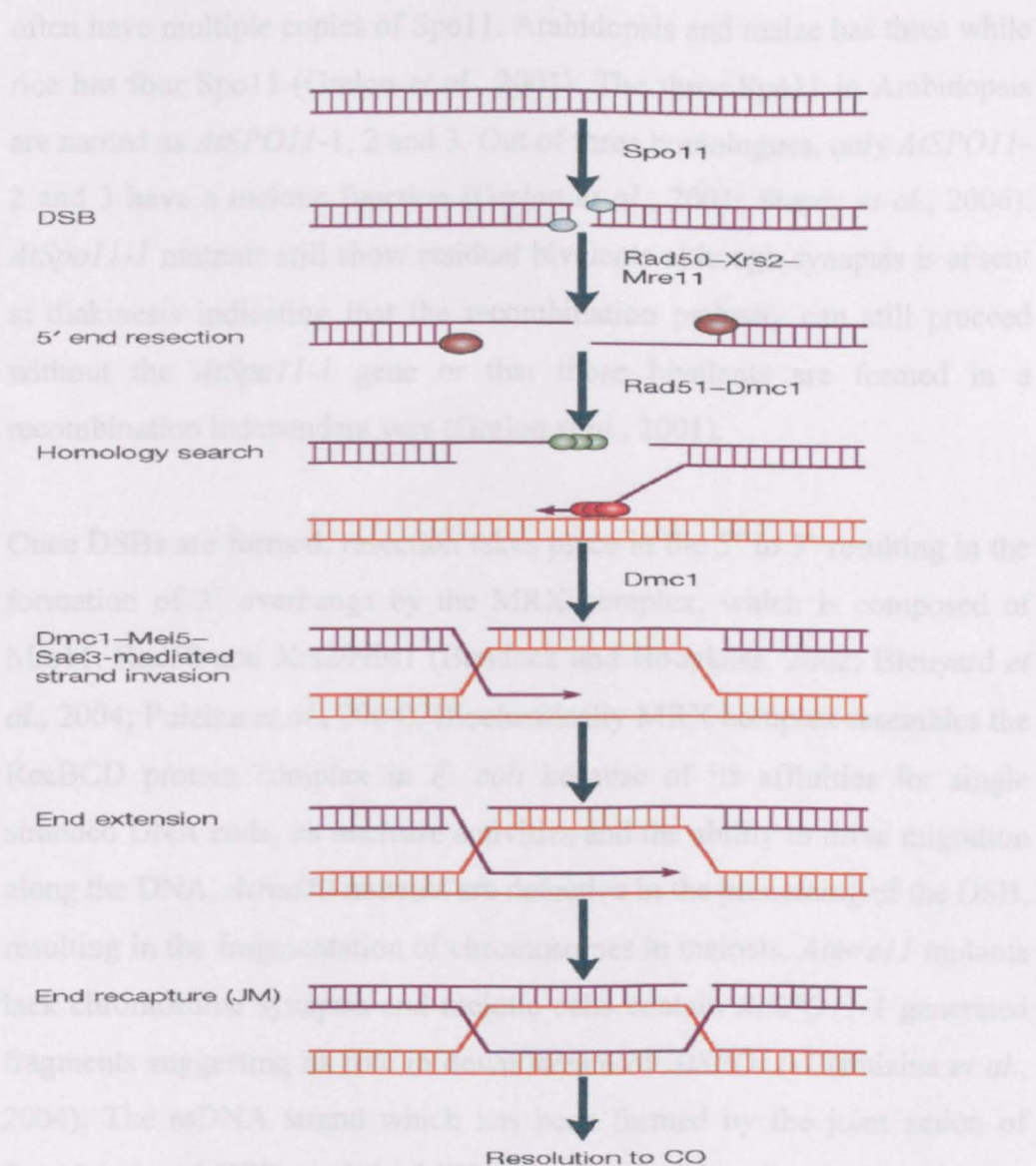
**Figure 1.9** Schematic representation of DSB induced HR in *E. coli*. Red and Blue bars represent homologous DNA strands. Pre-synapsis stages (a-c). Upon DSBs, the ends are nucleotically processed to result in 3' overhangs. The synapsis stage is indicated by (d) when the processed broken DNA forms a Holliday junction with the help of strand invasion and strand exchange activities of the RecA protein. Post synapsis stages (e-f) involve DNA synthesis, branch migration and Holliday junction resolution.

### 1.9.5 HR during meiosis in *Saccharomyces cerevisiae*

Much of the understanding of HR in eukaryotes has originated from the studies conducted in budding yeast (*Saccharomyces cerevisiae*) and many genes have been identified that have a function in HR. The process of HR in budding yeast is controlled by a number of proteins, including; SPO11, RAD50, MRE11, XRS2, MER1, MER2, MEI4, MRE2, REC102, REC104, REC114 (Paques and Haber, 1999) and is initiated by the SPO11 enzyme through a trans-esterification reaction. Homologues of SPO11 have been found in *Arabidopsis thaliana*, *Caenorhabditis elegans*, *D. melanogaster*, *Coprinus cinereus*, and mammals which suggest its function is well conserved among organisms (Hartung and Puchta, 2001). Mutants lacking SPO11 showed reduced recombination and synapsis, but still showed the presence of residual bivalents during meiosis which may indicate a DSB-independent mechanism of pairing, synapsis or recombination (Loidl *et al.*, 1994) or those sufficient DSBs exist to allow a reasonable level of recombination events. After generation of DSBs with the help of SPO11, RAD50 along with MRE11 and XRS2 genes may serve a crucial role in resection of chromosomal ends to produce 3' single strand termini in budding yeast (Ogawa and Johzuka, 1995; Chamankhah and Xiao, 1999). Apart from resectioning of chromosomal ends this complex also plays a role in a set of events such as HR, NHEJ, telomere maintenance, and removal of SPO11 from the 5' end of newly formed DSBs (Trujillo *et al.*, 1998).

After generation of the DSB and resectioning, 3' single strand DNA invasion is catalyzed by a group of proteins including *RAD51* and *DMC1* (Bannister and Schimenti, 2004; Aylon and Kupiec, 2004). This is later followed by DNA synthesis that extends the ends of the invading strand and then by recapture of this strand, which generates a joint molecule (JM) between parent molecules which is a prerequisite for synapsis and contains a double Holliday junction (Figure 1.10). The Holliday junction can then be resolved into a crossover or non-crossover event. The physical connection in the form of chiasmata between two homologous chromosomes held through a double Holliday junction is crucial for accurate homologue segregation at the first meiotic division (Gerton and Hawley, 2005).





**Figure 1.10** Schematic representation of HR in budding yeast. Duplex DNA is represented by ladders with rungs representing base pairs. The two coloured ladders represent different homologues. (Source: Gerton and Hawley, 2005).

### 1.9.6 HR in plants

Many genes have been isolated and characterized which are involved in HR in plants (Lichten, 2001; Keeney, 2001; Villeneuve and Hillers, 2001). Most of these genes have been characterized by a reverse genetics approach given the high degree of conservation of function among eukaryotes. In plants, recombination is initiated by the Spo11 protein as in other eukaryotes to form DSBs (Lichten, 2001; Keeney, 2001). Unlike other eukaryotes plants form DSBs (Lichten, 2001; Keeney, 2001). Unlike other eukaryotes plants



often have multiple copies of Spo11. Arabidopsis and maize has three while rice has four Spo11 (Grelon *et al.*, 2001). The three Spo11 in Arabidopsis are named as *AtSPO11-1*, 2 and 3. Out of three homologues, only *AtSPO11-2* and 3 have a meiotic function (Grelon *et al.*, 2001; Stacey *et al.*, 2006). *AtSpo11-1* mutants still show residual bivalents although synapsis is absent at diakinesis indicating that the recombination pathway can still proceed without the *AtSpo11-1* gene or that those bivalents are formed in a recombination independent way (Grelon *et al.*, 2001).

Once DSBs are formed, resection takes place in the 5' to 3' resulting in the formation of 3' overhangs by the MRX complex, which is composed of Mre11, Rad50 and Xrs2/Nbs1 (Bundock and Hooykaas, 2002; Bleuyard *et al.*, 2004; Puizina *et al.*, 2004). Biochemically MRX complex resembles the RecBCD protein complex in *E. coli* because of its affinities for single stranded DNA ends, its nuclease activities and the ability to drive migration along the DNA. *Atrad50* mutants are defective in the processing of the DSB, resulting in the fragmentation of chromosomes in meiosis. *Atmre11* mutants lack chromosome synapsis and meiotic cells contain *AtSPO11-1* generated fragments suggesting its role in down stream of *AtSPO11-1* (Puizina *et al.*, 2004). The ssDNA strand which has been formed by the joint action of Spo11-induced DSBs and the MRX complex then invades the homologous double stranded DNA. This step is catalyzed by the *AtRAD51* and *AtDMC1* gene products which are homologues of the *E. coli RecA* recombinase that helps in homology searching, strand invasion and possesses ssDNA binding activity and DNA-dependant ATPase activity (Anderson *et al.*, 1997; Bishop *et al.*, 1992). Both *RAD51* and *DMC1* homologues have been characterized in many species, including Arabidopsis, maize, lily and rice (Anderson *et al.*, 1997; Ding *et al.*, 2001; Doutriaux *et al.*, 1998; Franklin *et al.*, 1999; Klimyuk *et al.*, 1997; Li *et al.*, 2004; Pawlowski and Cande, 2005).

The strand invasion event can lead to one of two pathways: the double – strand break repair (DSBR) pathway or the synthesis-dependant strand-annealing (SDSA) pathway (Bishop and Zickler, 2004). More importantly

during the DSBR pathway the DNA synthesis from the invading strand ligation results in the formation of double-Holliday junction and resolution of the Holliday junction produces either crossovers or non-crossovers. In plants such as *Arabidopsis* and other vertebrates, *XRCC3* gene - a paralogue of *RAD51* plays an important role in the resolution of Holliday junctions (Thompson and Schild, 2001; Brenneman *et al.*, 2002; Liu *et al.*, 2004; Symington and Holloman, 2004).

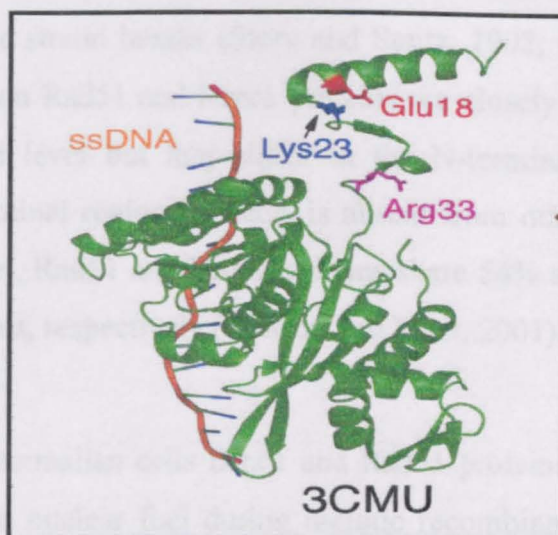
## **1.10 Recombinases**

### **1.10.1 *E. coli* RecA recombinase**

The *E. coli* RecA recombinase was one of the first proteins to be isolated that was reported to play a central role in the DNA recombination pathway (Clark and Margulies, 1965). RecA was first isolated in the 1960s in John Clark's lab during the screening of mutants that were defective in genetic recombination. RecA was found to exhibit an extreme sensitivity to both UV and X-rays. Biochemical studies reveal that RecA is in fact a DNA binding protein and promotes DNA-DNA interactions and electron microscopic observation of the RecA-DNA complexes led to the realization that these were nucleoprotein filaments, in which RecA formed a helical protein sheath around the DNA (DiCapua *et al.*, 1982). The DNA within the filament is composed of 6 RecA monomers per helical turn and the nucleoprotein filament elongates to approximately 1.5 times the length of the normal duplex DNA (Figure 1.11). The filaments provide a specialized architectural configuration for successful homology searching between two DNA molecules.

The RecA protein is a DNA-dependant ATPase that is able to bind ssDNA and promotes strand invasion and exchange between homologous DNA molecules (Ogawa *et al.*, 1979). In brief the functions of RecA during recombination in *E. coli* are four-fold. Firstly, it catalyzes the DNA interactions necessary for the establishment of homologous contacts. Secondly it promotes strand transfer between DNA molecules. Thirdly it provides a protein filament within which the DNA interactions take places

and finally it recruits more than 20 genes that function in the SOS response to DNA damage.



**Figure 1.11** 3D diagrammatic representation of *E. coli* RecA protein RecA-ssDNA-ADP-AlF<sub>4</sub><sup>-</sup> presynaptic filament (Chen *et al.*, 2008)

### 1.10.2 *RAD51* and *DMC1* recombinases

*RAD51* and *DMC1* are two homologues of the *E. coli* RecA gene (Thacker, 1999) which were first discovered in budding yeast (*Saccharomyces cerevisiae*). *ScRAD51* gene was isolated based on its sequence similarity with the prokaryotic *RecA* by Shinohara *et al.* (1992) and Aboussekhra *et al.*, (1992). Later Doug Bishop (1992) and Schwacha *et al.*, (1997) used a differential screening approach to isolate meiosis-specific transcripts in yeast. The gene they identified which was involved in meiotic recombination was found to be the *DMC1* gene which has a remarkable similarity to *RAD51* and *RecA*. In yeast *RAD51* gene is essential for repair of double strand breaks in mitotic cells. In meiosis both *rad51* and *dmc1* mutants have similar kinds of phenotype and both are essential for completion of the meiotic cycle and production of viable spores (Shinohara *et al.*, 1992; Bishop *et al.*, 1992). Homologues of *RAD51* and *DMC1* are found in mouse and humans. While full length *DMC1* is meiotically

expressed in yeast, in mouse a shortened form is expressed in somatic tissues (Shinohara *et al.*, 1993; Sato *et al.*, 1995; Habu *et al.*, 1996). Both proteins polymerize onto DNA to form nucleofilaments, promote strand exchange and are required for HR during meiosis and the recombinational repair of double strand breaks (Story and Steitz, 1992; Story *et al.*, 1992). Yeast and human Rad51 and Dmc1 proteins are closely related to RecA at the amino acid level but they differ in the N-terminal regions and the extended C-terminal region of RecA is absent from other proteins. At the amino acid level, Rad51 and Dmc1 proteins share 54% and 45% identity in humans and yeast, respectively (Masson and West, 2001).

In yeast and mammalian cells Dmc1 and Rad51 proteins have been shown to co-localize to nuclear foci during meiotic recombination indicating the combined role played by both the paralogues. The appearance of Rad51/Dmc1 foci coincides with the maxima of DSB and the majority of the foci seem to disappear soon after the chromosomes synapse. The co-localization studies are underpinned by the fact that both proteins are expressed simultaneously during the leptotene to pachytene stages of meiotic prophase I. They have also been shown to interact through two hybrid screening. It was initially unclear why eukaryotic cells contained two RecA homologues. The most plausible explanation for the presence of two RecA homologues is that both Rad51 and Dmc1 might have distinct and overlapping roles during recombination. This possibility was later proved by Schwacha and Kleckner (1997) by studying the role of Dmc1 and Rad51 in inter-homologue recombination. They used *ScRED1* mutants that are defective in the formation of the synaptonemal complex and prevent pathway-specific production of inter-homologue recombination products. The *Scdmc Scired1* double mutants were shown to contain DSB intermediates that recombine with a sister chromatid rather than with a homologue and the *Scred1 Scrad51* double mutants showed reduced levels of inter-homologue recombination. These results indicate that Dmc1 plays a meiosis-specific role, helping to promote recombination between homologous chromosomes whereas Rad51 permits inter-sister

recombination. So in yeast and other eukaryotes, Rad51 was shown to be involved in both mitosis and meiosis, but Dmc1 is a meiosis-specific Rad51-like protein (Bishop *et al.*, 1992; Shinohara *et al.*, 1992). This is because both Rad51 and Dmc1 are regulated by distinct sets of accessory factors. In the case of Dmc1, its ability to contribute to meiotic recombination is influenced by its interacting with accessory factor proteins including Mei5-Sae3, Hop2-Mnd1 and Tid1/Rdh54, even though some of these proteins also interact with Rad51. The main interacting accessory factor protein factors include Rad52, Rad55-Rad57 and Rad54. Schwacha and Kleckner (1997) using 2D gel analysis of *Scdmc1* and *Scrad51* mutants showed that the ScDmc1 protein specifically promotes interhomolog recombination and possesses an interhomolog-specific activity while ScRad51 lacks it. In Arabidopsis, the *atrad51* mutant is sterile and shows defects in pairing and synapsis during meiosis and also severe chromosome fragmentation due to failure to repair DSBs induced by Spo11 (Li, 2004), whereas the *atdmc1* mutant is not completely sterile and univalents instead of bivalents are observed at late prophase I (Couteau *et al.*, 1999) and no chromosome fragmentation is observed in the *atdmc1* mutant. These results suggest that the Dmc1 protein functions in the selective invasion of the homologous chromosome during meiosis whereas the repair of Spo11-induced DSBs are mediated by the Rad51 protein between sister chromatids by an alternate pathway.

### 1.11 *RAD51* gene

The *RAD51* gene is a eukaryotic homologue of *E. coli* RecA based on its function during HR such as homology searching, pairing and stand-exchange between two different DNA molecules. The phenotypic similarities between RecA and the Rad51 mutants also provided strong evidence that both are true homologues. *RAD51* homologues are found in fungi, animals, humans and plants. In yeast, the *RAD51* gene is induced after gamma irradiation (Shinohara *et al.*, 1993; Stassen *et al.*, 1997). In mice and humans, *RAD51* foci are found in cells undergoing meiotic recombination and the foci are associated with the chromosomes at the onset

of synapsis and synaptonemal complex formation (Daboussi *et al.*, 2002). Even though yeast *rad51* mutants are highly sensitive to DNA-damaging agents such as ionizing radiation and MMS but they are still viable. By contrast, homozygous *RAD51* knockout mice are embryo lethal which indicates an essential role played by *RAD51* in their development. Experiments conducted in chicken DT40 cell lines by Takata *et al.*, (2001) and colleagues resulted in the observation that the lack of Rad51 was accompanied by an increase of chromosomal breaks indicating the role of Rad51 in the maintenance and/or repair of replication forks. In yeast the role of Rad51 is not restricted to DNA repair at mitotic recombination and recombination intermediates formed at meiosis revealed that the *Scrad51* mutants accumulated DSBs and were defective in their conversion to joint molecules. In addition, other meiotic defects such as reduction in homologue pairing, delayed synapsis and a decrease in viability during sporulation were observed in yeast (Shinohara *et al.*, 1992; Rockmill *et al.* 1995).

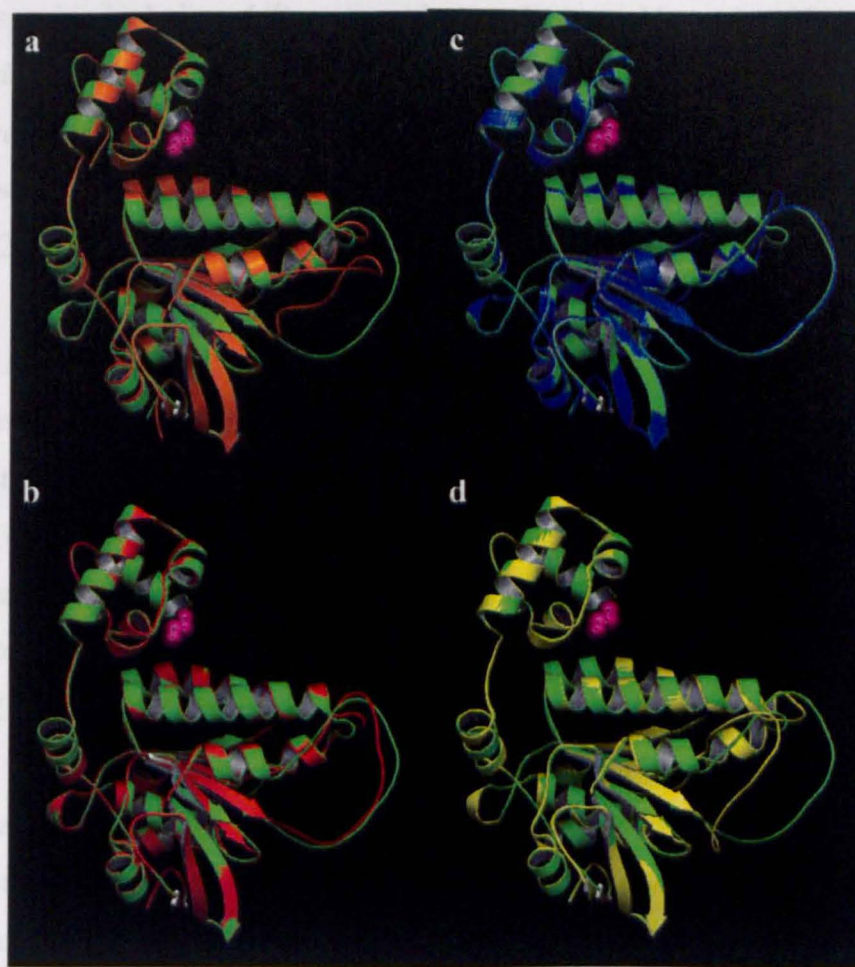
The homologues of *RAD51* have been cloned from several plant species (Terasawa *et al.*, 1995; Doutriaux *et al.*, 1998; Franklin *et al.*, 1999). Genes paralogous to *RAD51* have also been identified in Arabidopsis namely *AtRAD51B*, *AtRAD51C* and *AtRAD51D*. The *AtRAD51B* gene is mainly involved in somatic recombination and the *atrad51b* mutants produced viable seed indicating no essential role for *AtRAD51B* in meiosis (Osakabe *et al.*, 2002). By contrast the *AtRAD51C* gene is involved in normal meiotic chromosome synapsis and DSB repair of DNA (Li *et al.*, 2005). Analysis of *atrad51c* mutant plants suggests that it plays a vital role in both male and female gametogenesis but was not essential for mitosis (Bleuyard *et al.*, 2004; Li *et al.*, 2005). The *AtRAD51D* gene is believed to play a role in HR and it may also have a broader role in pathogenesis-related (PR) gene regulation during systemic acquired defense in plants (Durrant *et al.*, 2007). In *Zea mays* there are two closely related *RAD51* genes, namely *Zmrad51A1* and *Zmrad51A2* (Franklin *et al.*, 1999), which are yeast homologues of the *RAD51* gene and encode RecA-related proteins. Phylogenetic analysis indicates that these two genes are most closely related to Arabidopsis

*RAD51* and distantly related to the other *Arabidopsis* paralogues (*AtRAD51B*, *AtRAD51C*, *AtRAD51D*, *XRCC2* and *XRCC3*). Analysis of *Zmrad51A1* and *Zmrad51A2* single and double mutants in maize suggested that *Zmrad51A1* and *Zmrad51A2* have redundant functions. Maize plants with single mutations in one *rad51* gene survive and flower but *rad51* double mutants are highly sensitive to radiation treatment, male sterile and show a reduced production of viable female gametes. However, in the female gametes the actual rate of meiotic crossovers does not differ much to that of wild type, suggesting that reduced gamete viability is due to factors other than incorrect disjunction (Li *et al.*, 2007).

### 1.11.1 Structure of the Rad51 protein

The structure of the Rad51 protein can be broadly divided into three regions: an N-terminal domain, a linker region and a C-terminal domain. The linker region has a single  $\alpha$ -helix consisting 9 aa residues and the C-terminal domain contains the Walker A and Walker B motifs (Walker *et al.*, 1982). The two Walker motifs in the C-terminal domain give Rad51 both its ATP-binding and hydrolysis abilities allowing it to function as an ATPase (Miller *et al.*, 2004; Zhang *et al.*, 2005). Unlike the RecA protein, Rad51 protein has a modified N-terminal domain which house  $\alpha$ -helices to form a (HhH)<sub>2</sub> domain and binds DNA strand non-specifically (Shao and Grishin, 2000). Analysis of predicted secondary structure of *TaRAD51A1* by Khoo *et al.*, (2008) indicated that N-terminal of this protein has five  $\alpha$ -helices that are stacked upon each other to form two consecutive HhH motifs. A high level of structural conservation and functional similarity was observed by comparison of 3D models of *ScRAD51*, *AtRAD51*, *ZmRAD51A1* and *TaRAD51A1* predicted by using *in silico* molecular substitution by Khoo *et al.*, (2008) (Figure 1.12).





**Figure 1.12** 3D predictive protein modelling with ScRAD51 (a), HsRAD51 (b), AtRAD51 (c) and ZmRAD51A1 (d) superimposed onto TaRAD51A1. TaRAD51A1 is represented in green (A–D) (Source: Khoo *et al.*, 2008).

### 1.12 *DMC1* gene

*DMC1* is a homologue of the *E. coli* RecA protein and plays a central role in recombination, SC formation and cell-cycle progression in budding yeast (Bishop *et al.*, 1992; Bishop, 1994). *DMC1* gene homologues are found in a wide variety of organisms including fungi, mammals and plants (Masson and West, 2001). *DMC1* mutants in *S. cerevisiae* are defective in reciprocal recombination, fail to form normal SCs, and the cells are arrested in late meiotic prophase (Bishop *et al.*, 1992). By contrast, the disruption of mouse *DMC1* affects meiosis leading to cell apoptosis (Pittman *et al.*, 1998; Yoshida *et al.*, 1998). Homologues of *DMC1* have been identified in many



plant species including *Lilium longiflorum* (Kobayashi *et al.*, 1993, 1994), *Arabidopsis thaliana* (Sato *et al.*, 1995; Doutriaux, *et al.*, 1998), *Hordeum vulgare* (GenBank Accession Number AF234170), *Glycine max* (GenBank Accession Number U66836) and *Oryza sativa* (Kathiresan *et al.*, 2002; Ding *et al.*, 2001; Shimazu, *et al.*, 2001).

Yeast *DMC1* (*LIM15*) gene was isolated and found to hybridize with the meiosis-specific gene in *Lilium longiflorum* (Kobayashi *et al.*, 1993). Based on Southern and Northern blot analysis, *Arabidopsis* was found to contain a single copy of the *DMC1* gene which is expressed in young flower buds cells undergoing meiosis and in root tip cells (Sato *et al.*, 1995; Klimyuk and Jones, 1997; Doutriaux *et al.*, 1998). Unlike in yeast and mouse, *AtDMC1* is found to be expressed in vegetative tissues such as leaves at low levels and at very high levels in exponentially growing cells from suspension cultures (Klimyuk and Jones, 1997; Doutriaux *et al.*, 1998). Initial analysis of mutants of *AtDMC1* by Couteau *et al.*, (1999) suggested that plant *DMC1* homologues may play an important role in the pairing of homologous chromosomes during meiosis based on their observation that in *atdmc1* mutants aberrant chromosome behaviour in prophase I is observed. Such a failure to undergo meiotic synapsis and pairing of homologous chromosomes resulted in abnormal pollen grain formation and reduced fertility. The mutants however showed normal vegetative development. Homologues of *DMC1* were isolated from a commercially available genomic library of rice indica type strain IR36 and were termed *RiLIM15* by Sato *et al.*, (personal communication) and further characterization of the gene revealed that rice has two homologues of *RiLIM15* present in the genome. High levels of expression of the *RiLIM15* genes was observed in meiotic young panicles, cultured cells but not in mature leaves (Shimazu *et al.*, 2001). Kathiresan *et al.*, (2002) were able to isolate and establish that rice has two unlinked copies of the *DMC1* gene in rice located on chromosome 12 and 11, respectively and designated them *DMC1A* and *DMC1B*. Deng and Wang (2007) showed that that *OsDMC1* deficiency mediated by RNAi caused abnormalities in homologous pairing in rice as observed by FISH experiments and that these defects in turn lead to changes

in meiotic cell cycle progression, although they fail to block meiotic division in rice.

### 1.12.1 Dmcl protein structure

Unlike RecA and Rad51, human Dmcl formed nucleofilaments composed largely of stacked octameric rings on the DNA rather than helical structures (Masson *et al.*, 1999; Passy *et al.*, 1999). This was surprising because helical nucleoprotein filaments represent the biologically active form of RecA/Rad51. Even though no helical structure of filaments were observed for Dmcl, it was found that the stacked rings were in fact functional and promote homologous chromosome pairing and strand exchange in an alternate pathway involving a three-stranded nucleoprotein intermediate instead of a pathway involving the separation and reannealing of DNA strands (Gupta *et al.*, 2001). The N-terminal of Dmcl showed higher affinity for both ss- and ds-DNA while the C-terminal is the active site of homomultimer formation (Nara *et al.*, 2000). Smith and Nicolas (1998) suggested that the N-terminus may mediate interactions with other proteins during recombination.

### 1.13 RAD51 paralogues

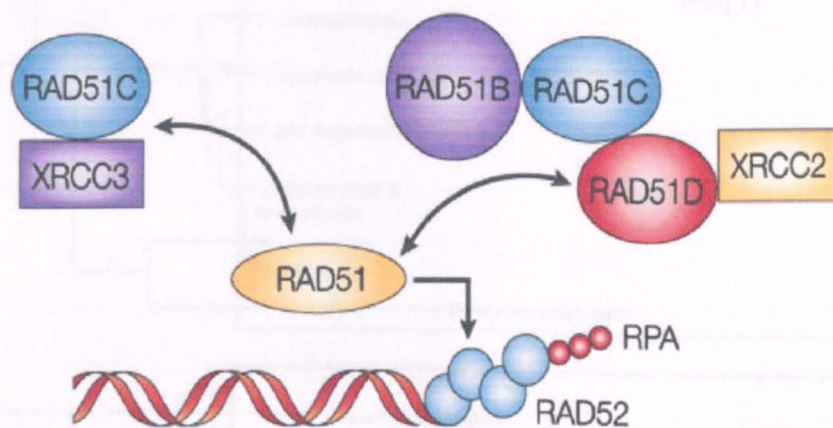
The finding that the *in vitro* activities of eukaryotic RAD51 are reduced considerably in comparison with bacterial RecA, led to the discovery of other protein cofactors which are involved in HR in eukaryotes. They were later identified in screens for x-ray-sensitive mutants in budding yeast. In yeast, RAD51 is stimulated by RAD51-like proteins: RAD52 (Sung 1997a; New *et al.*, 1998; Shinohara and Ogawa, 1998), RAD54 (Petukhova *et al.*, 1998, 1999; Mazin *et al.* 2000; Solinger *et al.* 2001), RAD55/57 (Sung, 1997b) and RP-A (Sung, 1994; Sigurdsson *et al.*, 2001). RAD52, RAD55 and RAD57 paralogues are important for DSB-induced HR in budding yeast. Within the RAD52 epistasis group, the sub-family of the RAD51-like proteins are of considerable important because of the finding that RAD51 is the homolog of bacterial recombinase RecA (Aboussekhra *et al.*, 1992;

Shinohara *et al.*, 1992. ScRad51 proteins bind to DNA in an ATP-dependant manner as does the bacterial RecA protein and also plays an important role in DSBs repair and/or HR (Dudas and Chovanec, 2004). Yeast *dmcl* mutants show meiotic defects (Bishop *et al.*, 1992), while Rad55 and Rad57 proteins are thought to play an important role in assembly or stabilization of multimeric forms of Rad51 and Rad52 and Replication Protein A (RPA) (Sung, 1997) and yeast mutants lacking Rad55 and Rad57 show an increase in DNA damage and sensitivity to lower temperatures. Over-expression of Rad51 suppresses this sensitivity (Lovett and Mortimer, 1987).

In 1980s, two mammalian cell lines (*irs1* and *irs1SF*) were identified and found to be sensitive to ionizing radiation, U.V. and cross-linking agents (Jones *et al.* 1987; Fuller and Painter, 1988). The genes responsible for this behaviour were then identified to be mammalian paralogues of *RAD51*: *XRCC2* and *XRCC3* (Tebbs *et al.*, 1995; Tambini *et al.*, 1997). Amino acid sequence analyses of the *XRCC2* and *XRCC3* genes showed that the protein products share significant sequence homology with the cellular recombinase *RAD51* and are most likely to have arisen through gene duplication and divergent evolution (Cartwright *et al.*, 1998b; Liu *et al.*, 1998). Later three further *RAD51* paralogues were identified by database analyses (Albala *et al.*, 1997; Rice *et al.*, 1997; Cartwright *et al.*, 1998a; Dosanjh *et al.*, 1998; Pittman *et al.*, 1998): *RAD51B* (also known as *RAD51L1*, *hREC2*, *R51H2*), *RAD51C* (*RAD51L2*) and *RAD51D* (*RAD51L3* or *R51H3*). Amino acid sequence analysis and functional relationships indicate that the vertebrate *RAD51* paralogues were closely related to yeast *RAD55/57* and play an important role in *RAD51*-mediated reactions. In mouse, individual gene disruption of *RAD51B*, *RAD51D*, or *XRCC2* leads to embryonic lethality and shows the paralogues are required for animal viability (Shu *et al.*, 1999; Deans *et al.*, 2000; Pittman and Schimenti, 2000). By contrast Chicken DT40 cells deficient in all five *RAD51* paralogues show reduced growth rates, chromosomal irregularities, and accumulation of spontaneous breaks due to inability of the cells to repair broken replication forks, but this does not lead to cell death (Takata *et al.*, 2000, 2001).

The genome of *Arabidopsis* was also found to encode seven Rad51-like proteins. Rad51, Dmc1, Rad51C and Xrcc3 were identified and cloned based on their homologies to yeast or human Rad51-like sequences. (Doutriaux *et al.*, 1998; Klimyuk and Jones, 1997; Osakabe *et al.*, 2002; Sato *et al.*, 1995; Urban *et al.*, 1996) and later phylogenetic tree analysis has allowed the identification of *Arabidopsis* homologues of Rad51B, Rad51D and XRCC2 (Bleuyard and White, 2004). *AtRAD51* is not essential for maintenance of genome integrity under normal conditions as *atrad51-1* mutants grow normally vegetatively but are completely male and female sterile (Li *et al.*, 2004). *Arabidopsis* mutants lacking Rad51B (AtRad51B), Rad51C (AtRad51C) and Xrcc2 (AtXrcc2) proteins confer hypersensitivity to the DNA cross linking agent MMC but not to ionizing radiation and only mutants defective in the *AtRAD51C* and *AtXRCC3* genes show meiotic defects, indicating the role played by these Rad51 paralogues in DNA repair is conserved across eukaryotes and only AtRad51C and AtXrcc3 (which together form the CX3 complex) play an essential role in meiosis in *Arabidopsis* and other eukaryotes. Recently Khoo *et al.*, (2008) reported the isolation and preliminary characterization of four *RAD51* paralogues in bread wheat based on known *RAD51* sequences of other species such as rice and other cereals: *TaRAD51A1*, *TaRAD51A2*, *TaRAD51C* and *TaRAD51D*.

Yeast Two-hybrid and immunoprecipitation experiments have shown that in eukaryotes, the five Rad51 paralogues form two complexes: an heterodimer complex composed of Rad51C and Xrcc3 and an heterotetramer complex composed of Rad51B, Rad51C, Rad51D and Xrcc2 (Liu *et al.*, 2002; Masson *et al.*, 2001; Miller *et al.*, 2002, 2004; Schild *et al.*, 2000; Wiese *et al.*, 2002) (Figure 1.13). The sub-complex composed of Rad51B and Rad51C plays a mediator role in the assembly of the Rad51-ssDNA nucleoprotein filament in the presence of RPA similar to that of Rad55-Rad57 heterodimer in yeast (Lio *et al.*, 2003; Sigurdsson *et al.*, 2001; Sung, 1997). The Rad51C protein, the CX3 complex and a Rad51D/Xrcc2 heterotetramer complex is believed to be involved in homologous pairing (Kurumizaka *et al.*, 2001, 2002; Lio *et al.*, 2003).



**Figure 1.13** Interplay of different Rad51 paralogues during HR in yeast  
(Source: Masson *et al.*, 2001)

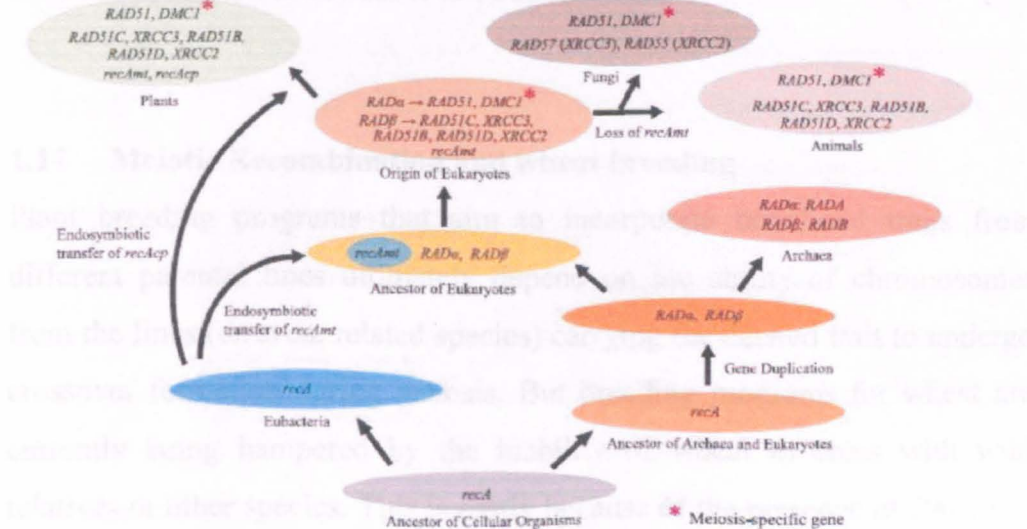
#### 1.14 Phylogenetic analysis of Rad51 and Dmc1 proteins

Phylogenetic analysis of yeast, human, mammalian, plant Rad51 and Dmc1 amino acid sequences was done by Kathiresan *et al.*, (2002). The phylogenetic tree showed that clear demarcation existing between the two homologues of the prokaryotic RecA protein and clearly sets Dmc1 apart from Rad51. However there is a close association between the two groups which indicates that both the genes are paralogous descendants of an ancient recombinase. While plant Dmc1 sequences cluster together the other eukaryotic Dmc1 sequences formed two further sub-clusters (Figure 1.14).



ancient *RecA* gene should have occurred before the divergence of Archaea and Eukaryotes to give rise to two lineages of *RAD51*-like genes, *RADα* and *RADβ*, but that the *RecA*-like gene has been maintained as a single copy gene in Eubacteria (except for few species). In Arachea, *RADA* and *RADB* were maintained as single-copy genes after the divergence from Eukaryotes with a subsequent loss of *RADB* in some cases. In the Eukaryote lineages further gene duplications in *RADα* and *RADβ* occurred before the divergence of the major eukaryotic groups. Gene duplications resulted in the formation of the *RAD51* and *DMC1* genes from *RADα* and *RAD51C*, *XRCC3*, *RAD51B*, *RAD51D* and *XRCC2* genes from *RADβ*, respectively (Figure 1.15). Some genes were lost during the evolution of species, for example the *DMC1* gene was lost from some insect and nematode species and *RADβ* was lost from several fungal and invertebrate lineages. The mitochondrion-derived *RecA* gene has undergone further duplications in flowering plants but was lost in the ancestor of animals and fungi.

are defective in specifying homologous recombination (Kobayashi and Ohtsuka, 1997; Anger et al., 1992; Franklin and Gubina, 2002; Pardo et al., 2002; Pardo et al., 2002).



**Figure 1.15** A model for the evolutionary history of *RecA/RAD51* recombination genes (Source: Lin *et al.*, 2006)



### 1.16 Other meiotic recombination genes

In *E. coli*, three related proteins called MutS, MutL and MutH were reported to play an important role in repair of DNA mismatches (Modrich and Lahue, 1996). These Mismatch Repair genes are highly conserved, but in higher eukaryotes the three-protein bacterial systems have been replaced with MutS and MutL with additional protein accessories (Kolodner, 1996). Apart from their role in mismatch repair, they also play an important role in meiotic recombination. The meiotic role played by MMS repair proteins has been described in detail in budding yeast. Meiotic genes homologous to those in other organisms have been isolated from plants like *Arabidopsis thaliana* (Klimyuk and Jones, 1997; Sato *et al.*, 1995), lily (Kobayashi *et al.*, 1994) and wheat (Dong *et al.*, 2002; Ji and Langridge, 1994). Most of the genes involved in DSB repair and meiotic recombination in plants has been characterized by reverse genetics approaches. The classic example is the discovery of *desynaptic1*, *desynaptic2*, and *phs1* genes in maize mutants that are defective in specifying homolog recognition (Golubovskaya and Grebennikova, 1997; Auger *et al.*, 1997; Franklin and Golubovskaya, 2003; Bass *et al.*, 2003; Pawlowski *et al.*, 2004).

### 1.17 Meiotic Recombination and wheat breeding

Plant breeding programs that aim to incorporate beneficial traits from different parental lines ultimately depend on the ability of chromosomes from the lines (or even, related species) carrying the desired trait to undergo crossover formation during meiosis. But breeding programs for wheat are currently being hampered by the inability of wheat to cross with wild relatives or other species. This is partly because of the presence of *Phl* gene which restricts homoeologous (non-homologous) chromosome pairing and promotes homologous chromosome pairing. Complete deletion or constitutive silencing of *Phl* gene can have deleterious effects on the chromosome pairing during meiosis. Al-kaff *et al.*, (2007) observed that the deletion of 5B *cdk-like* (of which *Phl* gene is a part of ) locus results in activation of the 5A and 5D *cdk-like* loci. A potential solution could be to switch off the *Phl* in hybrid plants between wheat and the wild relative and



switch it on again in  $F_1$  plants so that stable recombinant lines will be formed. As the *Ph1* locus is found to be similar to CDK2, it may be possible to identify or design drugs that would be delivered directly into flowers of the hybrid plants, thereby inhibiting *Ph1* function during the meiosis of the hybrids. Many compounds were found to inhibit and inactivate CDK2 function in mammals and yeast (Shapiro, 2006) but many of them can have deleterious effects and can cause plant cell death because they can target other CDK2 proteins as well. So an alternative solution would be to find natural meiotic mutants in wheat which allow non homologous chromosome pairing in wheat with its relatives.

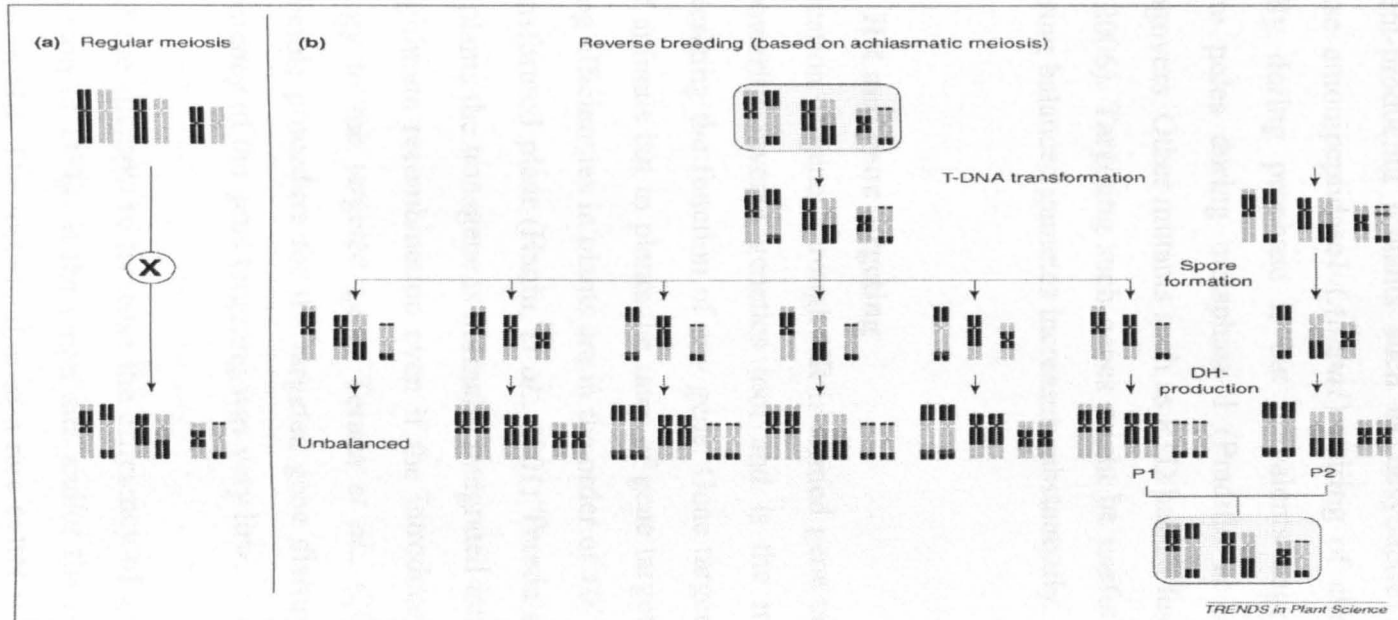
Recently three kinds of natural meiotic mutants have been found in plants that promote non-homologous chromosome pairing. The first one is the ZYP1a/ZYP1b double mutant in *Arabidopsis* (Higgins *et al.*, 2005). ZYP1a and ZYP1b are redundant components of the central region of the SC complex. The absence of both the genes results in 80% reduction of crossover frequency but more importantly some crossovers are formed between non-homologous chromosomes. Orthologues of ZYP1 were found to be present in Rye (Mikhailova *et al.*, 2006) and *Triticum monococcum* (Osman *et al.*, 2006). The second report of crossover formation in homologous chromosomes is from a *RAD51* mutant in maize (Li *et al.*, 2004). While *Arabidopsis* plants lacking *RAD51* are completely sterile (Li *et al.*, 2004), maize plants lacking *RAD51* are only fully male sterile and the surviving female gametes carry a normal numbers of crossovers. Cytologically at Metaphase I, multivalents and bivalents are observed that were composed of non-homologous chromosomes. The final report of crossover formation between non-homologous chromosomes is of an uncloned *sy10* mutant in rye where cytologically non-homologous chromosomes are observed to pair (Mikhailova *et al.*, 2006). None of the above genes are yet characterized in detail in wheat, even though Khoo *et al.*, (2008) recently published a report of the *RAD51* gene family in bread wheat and their conservation across eukaryotes.

The frequency and distribution of recombination events during meiosis could have a major effect in wheat breeding programs. Any increase of these frequencies and altering the distribution of recombinational events could have profound effects, reducing linkage drag and increasing the number of breakpoints per meiosis. This has value for the localization of genes prior to cloning by increasing genetic distances and potentially it may also improve gene transfer from alien species in wheat. Over-expression or silencing of many genes involved in crossover formation has found to have an effect on recombination frequencies (Rozwadowski, K.L. and Lydiate, D.JWO/2002/022811). Masson and Paszkowski, (1997) reported an uncharacterized *xrs4* mutant of *Arabidopsis* in which recombination frequency in certain regions increased by two-fold. In tomato, over-expression of the *MLH1*, a mismatch repair protein, led to 10% increase in chiasma frequency (Wittich *et al*, WO/2007/030014). Betzner *et al.*, (WO/2002/008432) observed a two-fold increase of recombination frequency on an *Arabidopsis* genetic interval by over-expression of the *RAD51* gene.

An example of one way in which manipulation of recombination genes to aid plant breeding has been suggested in Reverse Breeding by Siaud *et al.*, (2004) and Higgins *et al.*, (2004). They suggested that RNAi silencing of essential meiotic recombination genes such as *DMC1* can lead to total suppression of crossover formation in *Arabidopsis* has led to the concept of Reverse Breeding as an alternative to conventional plant breeding. Conventional plant breeding strategies attempt to combine valuable traits from different cultivars into new elite varieties, so they start with the selection of two homozygous lines with desired traits and sexually cross them to produce a novel elite variety combining the desirable traits from both the parents into a new hybrid. But a successful plant breeding program depends on the ability of plant breeders to bring together desirable traits into a new hybrid by combining the desired combination of alleles on chromosomes. Meiotic recombination has played a very important role in successful plant breeding because during meiotic crossing over takes place which results in the reshuffling of chromosome segments leading to novel

combination of alleles. However meiotic recombination also produces undesired effects during crossing-over including splitting up already good combinations of alleles.

Reverse Breeding does not start with the selection of homozygous parents for crossing but instead starts with a heterozygote (hybrid). The process of Reverse Breeding starts with the transformation of any elite heterozygote ( $F_1$ ) using a RNAi approach or a knockout technology that targets a gene that plays a role in crossover formation, such as *DMC1*. Total suppression of crossover formation leads to the production of low numbers of viable and also large numbers of unbalanced haploid gametes containing one, two or more copies of chromosomes because the homologues are not joined by chiasmata during prophase of Meiosis I and consequently they remain as univalents during anaphase I. These univalents subsequently segregate randomly during meiosis I to daughter cells. The balanced gametes containing one copy of each chromosome will be formed at a probability of  $(1/2)^x$ , where x is the basic chromosome number. Consequently, the chance of getting balanced gametes decreases exponentially with the chromosome number and for species in which the chromosome number equals 12 or greater, the approach seems infeasible (Dirks *et al.*, WO/2003/017753). The resulting plant produce a very low number of viable and balanced haploid gametes that are then regenerated into a DH plant and are completely homozygous. Among the doubled haploids, parents with complementary genotypes can be recruited that, on crossing, they will reconstitute the exact genotype of the elite hybrid again (Fig 1.16).



**Figure 1.16** Schematic representation of normal meiosis (a) and Reverse Breeding strategy (b)  
(Source: Wijnker and Jong, 2008).

A variety of mutants that lack crossovers and are univalent-producing have been identified in *Arabidopsis* and other plant species. In *Arabidopsis*, mutants lacking crossovers (almost) exclusively produce univalents (Ma, 2006; Mercier and Grelon, 2008), although their chromosome behaviour during meiotic prophase was different from that reported above. In some univalent-producing mutants such as *desynaptic1* (*Atdsyl*) and meiotic prophase *amonipeptidase1* (*Atmpal*) pairing of chromosomes takes place normally during prophase I, but univalents segregate preferentially to opposite poles during metaphase I (Pradillo *et al.*, 2007) without any crossovers. Other mutants such as PTD have a few crossovers (Wijeratne *et al.*, 2006). Targeting such genes might be useful because the chance of recovering balanced gametes increases substantially.

### 1.18 HR and gene targeting

Modification of genes through HR is termed gene targeting. Gene targeting is a powerful reverse genetics tool and is the most direct method for characterizing the function of any gene. Gene targeting is successful in the case of animals but in plants the rates of gene targeting are very low. Gene targeting efficiencies in plants are in the order of  $10^{-3}$  to  $10^{-4}$  targeted events per transformed plant (Hanin, *et al.*, 2001; Terada *et al.*, 2002) because in higher plants the transgene is normally integrated into the genome randomly by illegitimate recombination even if the introduced sequences have high homology to the targeted gene. Terada *et al.*, (2002) have developed a reproducible procedure for the targeted gene disruption by HR in rice but the efficiency of the gene targeting was very low.

Most of the attempts to increase the efficiency of gene targeting aim at the introduction of DSBs at the target site and/or the modification of proteins involved in HR. The introduction of a rare-cutter restriction enzyme site at target site increased the targeting efficiency to such a site by two-fold in the somatic tissues in which gene targeting is common (Puchta, *et al.*, 1996). But the DSB has to be introduced close to the gene and this creates a problem. Shaked *et al.*, (2005) reported an increase of between one and two

orders of magnitude in the GT frequency in *Arabidopsis thaliana* plants by over-expressing the yeast chromatin remodeling protein RAD54.

### 1.19 General project aims

Meiosis is a fundamental process occurring in all sexually reproducing species and meiotic recombination is the most important processes occurring during meiosis. Very little information was available on meiotic recombination and the genes controlling this in hexaploid wheat, principally due to its large genome size (~17,000 MB) and the polyploid nature of its genome. A detailed understanding of the genes and their gene homoeologues involved in meiotic recombination could aid in the development of strategies for increasing the frequency and altering the distribution of recombination events by inactivating or over-expressing recombination genes. This could reduce linkage drag and increasing the number of breakpoints per meiosis might aid in Reverse Breeding of the wheat crop, if this is a viable approach

The two main meiotic recombination genes involved in HR are *RAD51* and *DMC1*. Both the genes are RecA-like proteins and play an important role in binding of DNA, homology search, strand invasion and strand exchange resulting in the formation of Holliday junctions, which can be resolved to yield crossover or non-crossover events. Even though both the genes are well characterized in a wide variety of organisms including plants, very little information is available in wheat. In most of the diploid plant species, deletion of either *RAD51* or *DMC1* orthologues usually leads to sterility. Wheat, as a polyploidy, offers a unique opportunity to examine the effects of the deletion of specific homoeologues, while maintaining a level of fertility. Recent reports of the *RAD51* gene family in bread wheat by Khoo *et al.*, (2008) describe the functional conservation of this gene family across eukaryotes but it fails to characterize the role and function of individual homoeologues of the *RAD51* and *DMC1* genes in meiotic recombination in bread wheat. A preliminary study has been initiated with the objective of studying the *RAD51* and *DMC1* gene homoeologues in wheat and their

detailed characterization to study the role of individual gene homoeologues. The current research hypotheses are as follows:

- The predicted three homoeologues of each of the genes under study (*RAD51* and *DMC1*) are genes involved in meiosis and play an important and essential role in meiotic recombination based on functional conservation of these genes across eukaryotes.
- As these genes are involved in meiotic recombination any loss or deletion of each of the gene homoeologues would result in partial sterility. All three homoeologues of a particular gene are equally expressed and the loss or deletion of an individual copy of any of the genes could possibly be compensated either by over-expression or by the residual expression of the remaining copies.
- As these genes are homologues of *AtRAD51* and *AtDMC1* any loss of function of their mutants in *Arabidopsis* T-DNA lines can be compensated by the appropriate wheat *RAD51* or *DMC1* gene.

The main aim of my PhD is to study the *TaRAD51A*, *TaRAD51B* and *TaRAD51D*; *TaDMC1A*, *TaDMC1B* and *TaDMC1D* gene homoeologues to try to understand the role and function of individual homoeologues in meiotic recombination in wheat, using a mutant population of wheat and through an investigation of phenotype, genotype, cytology and gene expression of wild-type and mutant lines. The results will be compared to the function of *DMC1* and *RAD51* in *Arabidopsis thaliana* as already reported by Couteau *et al.*, (1999) and Li *et al.*, (2004) respectively, and through attempts to functionally complement the above genes in *AtDMC1* and *AtRAD51* T-DNA insertional inactivation lines. As such the research objectives of my PhD are detailed as below:

1. Microarray transcript expression profiling will be done for *RAD51* and *DMC1* genes of Arabidopsis, rice and wheat to confirm the earlier results and also to plan experiments accordingly (Chapter 3).
2. As the sequences of *RAD51* and *DMC1* genes (both genomic and cDNA) were not available in public databases, these will first be cloned and sequenced. These will be analyzed in detailed using bioinformatics tools such as multiple sequence alignments, phylogenetic analysis and proteins structure modelling. This will be a first step towards functional characterization of these homoeologous genes *in planta* (Chapter 4).
3. Genome specific tests will be designed to be able to distinguish individual homoeologues of *TaRAD51A*, *TaRAD51B* and *TaRAD51D*; *TaDMC1A*, *TaDMC1B* and *TaDMC1D* at both cDNA and gDNA level. These tests are important tools to characterize both *in planta* effects of these gene homoeologues in wheat and also for the functional characterization. *In planta* effects of the gene homoeologues in wheat will be studied by both reverse and forward genetics approach using a fast neutron and gamma ray irradiated population of wheat to identify mutants lacking specific homoeologues of *RAD51* and *DMC1* genes identified using the genome specific tests which have already been designed. The identified mutants will be characterized phenotypically and cytogenetically (Chapter 5).
4. Functional characterization of the *TaRAD51* and *TaDMC1* homoeologue genes will be undertaken by using the identified mutants along with comparisons with wild type wheat using Q-PCR to evaluate the level of expression of different homoeologues and explore whether there are compensatory changes in gene expression in the genome-specific deletion lines. Where available *TaRAD51* and *TaDMC1* gene homoeologue expression will be examined using



available microarray data of both *Arabidopsis* and wheat and the results correlated with Q-PCR data (Chapter 2 and 5).

5. Assessment of the effects of over-expression and any potential functional complementation of the *TaRAD51* and *TaDMC1* cDNA homoeologues will be done using representative homoeologues in *AtRAD51* and *AtDMC1* T-DNA lines respectively (Chapter 6).

## CHAPTER 2: MATERIALS AND METHODS

### 2.1 Materials

#### 2.1.1 List of chemicals used

Chemical Name	Company
ABgene PCR plates	Thermoscientific, UK
Agar Agar powder	Fisher Co. Ltd, UK
Agarose Electrophoresis grade	Invitrogen, UK
Agarose type PGP	Park Scientific Ltd, UK
Ampicillin sodium salt	Sigma-Aldrich Co. Ltd., UK
Boric acid	Melford, UK
Bromophenol blue	Sigma-Aldrich Co. Ltd., UK
Chloral hydrate	Sigma-Aldrich Co. Ltd., UK
Chloroform	Fisher Co. Ltd, UK
Dimethyl formamide (DMF)	Fisher Co. Ltd, UK
Ethanol Absolute	Fisher Co. Ltd, UK
Ficoll Type 4000	Fisher Co. Ltd, UK
Fuchsin acid	Fisher Co. Ltd, UK
Glacial Acetic acid	Fisher Co. Ltd, UK
Glucose	Fisher Co. Ltd, UK
Glycerol	Fisher Co. Ltd, UK
Hydrochloric acid	Fisher Co. Ltd, UK
Iodine crystals	Fisher Co. Ltd, UK

IPTG (Isopropylthio-b-D-galactoside)	Apollo
Kanamycin sulfate	Sigma-Aldrich Co. Ltd., UK
Magnesium Chloride	Fisher Co. Ltd, UK
Malachite Green Oxalate salt	Sigma-Aldrich Co. Ltd., UK
MS (Murashige and Skoog medium)	Duchefa Biochemie, Netherland
Orange G	Sigma-Aldrich Co. Ltd., UK
Orcein	Sigma-Aldrich Co. Ltd., UK
Phenol	Sigma-Aldrich Co. Ltd., UK
Potassium Chloride	Fisher Co. Ltd, UK
Potassium Iodide	Fisher Co. Ltd, UK
Propan-2-ol	Fisher Co. Ltd, UK
Rifampicin	Sigma-Aldrich Co. Ltd., UK
Sharp strip 8 strip PCR tubes	Anachem, UK
Sodium Chloride	Fisher Co. Ltd, UK
Sucrose	Fisher Co. Ltd, UK
Tris base (Tris (hydroxymethyl)aminomethane)	Melford, UK
Tryptone	Oxoid, UK
X-GAL (5-Bromo-4-chloro-3-indolyl $\beta$ -D-galactopyranoside)	Apollo, UK
Xylene cyanol FF	Fisher Co. Ltd, UK
Yeast Extract granulate	MERCK, Germany
Selwet L-77	Lehle seeds, USA
0.22 $\mu$ m filter	Anachem, UK

### 2.1.2 List of Equipment & Apparatus used

<b>Equipment/Apparatus</b>	<b>Company</b>
28°C incubator	Gallenkamp,UK
37°C incubator	Gallenkamp,UK
50 ml tube centrifuge; Centaur 2	Shelton Technical Ltd., UK
Binocular microscope; Vickers	Vickers Ltd., England
BioRad MicroPulser™	Bio-Rad Lab, UK
Cold Centrifuge; Centrifuge 5810R	Eppendorf, Germany
Digital camera with Binocular microscope; Infinity 2-2C	Lumenera corporation, Canada
Digital camera with Stereo microscope; Leica DC480	Leica Microsystem Ltd., UK
Gel-electrophoresis power supply; PowerPac Basic™	Biorad lab, UK
Gel-electrophoresis tanks; Sub-cell®GT and Mini-Sub®Cell GT	Biorad lab, UK
Gene Amp PCR system 9700(PCR)	Applied Biosystems
Hotblock- DRI-Block DB-3	Technique, UK
Large Autoclave 400/500L	Boxer Laboratory Equipment Ltd. UK
Large orbital incubator; Gallenkamp, 10x400.xx1.0	Sanyo-biomedical, UK
Leica GLS 50X microscope	Leica
LightCycler® 480 instrument	Roche Applied Science, UK
Microcentaur	MSE
Microcentrifuge; Sigma 1-15, 10115	Sigma, Germany
MX3005p real-time PCR machine	Stratagene
Nanodrop (ND-100)	Thermo Scientific
Omega media prestige medical	Prestige™Medical, England

pH meter; SevenEasy pH	Mettler Toledo, UK
Small orbital incubator; Stuart ,S150	Bibby Sterilin Ltd., UK
Spectrophotometer; SP6-500 UV	Pye Unicam Ltd., England
Stereo microscope; Leica MZ6 and Leica MZ75	Leica Microsystem Ltd., UK
Thermo Hybaid PCR machine	Thermo Electron corporation
UV image Capture-Gel Doc 2000	Biorad lab, UK
Video copy processor p91	Mitsubishi
Vortex machine; Genie-2	Scientific industries Inc., USA
Water bath (Grant JB series)	Grant Instruments (Cambridge), UK
Weight 2 digits; Precisa Junior 500C and Precisa XB4200C	Precisa Instrument Ltd., Switzerland
Weight 4 digits; Precisa XB120A	Precisa Instrument Ltd., Switzerland

### 2.1.3 Oligonucleotides

Oligonucleotides used for PCR were mainly obtained from commercial available suppliers Sigma-Aldrich and MWG Eurofins. All oligos were resuspended in TE Buffer (pH 8.0) to a final concentration of 200  $\mu$ M and the primers were diluted to a working concentration of 20  $\mu$ M with nuclease free water.

### 2.1.4 Computer analysis software used

**2.1.4.1 Accelrys Gene Version 2.0, Accelrys Software Inc., USA:** All primer design, multiple alignments of nucleotide and amino acid sequence were done using this software (<http://accelrys.com/>)

**2.1.4.2 EBI Clustal W2 Version 2.0:** All comparative amino acid sequence analysis was done using this server (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>).

- 2.1.4.3 MEGA (Multiple Evolutionary Genomic Analysis) 4.1 (beta), Center for Evolutionary Functional Genomics, USA:** The alignments of amino acid sequences and subsequent construction of the phylogenetic tree (neighbour joining method) (Saitou and Nei 1987) was carried out using this software. All the parameters used in the software are default except for the following: the pair-wise deletion option was used for the amino acid sequences, bootstrap values were fixed at 10,000 replicates and the model, setting was amino acid: Poisson correction with predicted gamma parameters set at 2.0 (<http://www.megasoftware.net/>)
- 2.1.4.4 Swiss Model server and Swiss pdb viewer:** The amino acid sequences were processed using Swiss model workspace (<http://swissmodel.expasy.org/workspace>, version 8.05, 2009). The computed outputs were then used to generate the 3D protein structures using Swiss-Pdb Viewer DeepView (<http://spdbv.vital-it.ch>, version 4.0, 2009).
- 2.1.4.5 SPSS statistical package (SPSS, version 16.0, SPSS Inc., USA):** The data from the experiments were analyzed by one-way analysis of variance (ANOVA) and differences among mean values were determined by the Duncan's multiple range test ( $P \leq 0.05$ ). All values were expressed as mean  $\pm$  standard error of mean (S.E.M.).
- 2.1.4.6 The Leica confocal Software (LCS):** The Leica TCS SP2 AOBS confocal scanning microscope was used to image the plant materials expressing fluorescence proteins in Arabidopsis pollen. The microscope is equipped with the Leica confocal software (LCS), a 100 mW multi-line Argon laser (458 nm, 476 nm, 488 nm, 496 nm and 514 nm) and a 1 mW He-Ne laser (543 nm) as excitation sources.
- 2.1.4.7 Quantity One Version 4.6.5, BioRad, UK:** Agarose gel images were viewed and analyzed using this software.

**2.1.4.8 Infinity Capture, Lumenera Corporation, Canada:** An advanced image capture and processing software package was used for imaging pollen and the meiotic stages of wheat and Arabidopsis.

**2.1.4.9 Leica IM50 version 4.0, Leica Microsystems Ltd., UK:** Image processing system was used for imaging wheat and Arabidopsis flower structure.

**2.1.4.10 Genespring GX10 software, Agilent technologies, USA:** Gene expression analysis software used for wheat meiotic gene expression profiling (<http://www.chem.agilent.com>).

**2.1.4.11 SIFT (Sorting Intolerant From Tolerant):** Is a program that predicts whether an amino acid substitution affects protein function based on protein homology and physical properties of the amino acids (<http://blocks.fhcrc.org/sift/SIFT.html>).

## **2.1.5 Plant materials used and growth conditions**

Wild type wheat stocks (*Triticum aestivum* c.v Highbury, Chinese spring and Paragon), Nullisomics-Tetrasomic stock (Sears, 1965) and mutagenized population of wheat (Fast-neutron irradiated population of Highbury wheat and Gamma-ray irradiated population of Paragon wheat) used in this work were obtained from John Innes Centre (Norwich). The wheat plants were grown both under field conditions and glasshouse settings. Under glasshouse conditions, the wheat stocks were grown in 13 cm pots using John Innes Centre compost 3.

T-DNA lines of Arabidopsis ecotype Wassilewskaja and Columbia were kindly provided by Dr. Marie-Pascale Doutriaux. Fluorescent tetrad lines (I<sub>1a</sub> and I<sub>3a</sub>) were kindly provided by Gregory P Copenhaver (University of North Carolina). The plants were grown in 9 cm pots using John Innes

Centre 3 compost mixed with 20 mg/l intercept (Scoott's, UK) under a 22 h photoperiod at 21-23°C under glasshouse conditions.

#### **2.1.6 Bacterial strains used**

The *Escherchia coli* strain DH5 $\alpha$  (Invitrogen, UK) was used for general manipulation of recombinant plasmids. *Agrobacterium tumefaciens* strain EHA105 was used for transformation of plant meiotic genes into *Arabidopsis*. Both the bacterial strains were obtained from Plant & Crop Sciences Division, Nottingham University, UK.

#### **2.1.7 Cloning vectors used**

The plasmid vector pGEM T-Easy (Promega, UK) (provided in the kit) was used for general cloning and sub-cloning purposes and the plasmid vector pGVT17 (kindly provided by Dr. Philip Vain, JIC, UK) was used for cloning wheat meiotic genes.



## 2.2 Methods

### 2.2.1 Bacterial culture growth media

Sterilized by autoclaving at 121°C for a 15 minute cycle unless otherwise stated

#### 2.2.1.1 LB (Luria-Bertani) liquid medium

10 g tryptone, 5 g yeast extract and 10 g NaCl was mixed with distilled water until dissolved completely. The pH was adjusted to 7.5 with KOH and the final volume was made up to 1 liter with distilled water. The media was sterilized by autoclaving

#### 2.2.1.2 LB Agar

As for LB liquid medium, but 15 g of Agar was added prior to autoclaving

#### 2.2.1.3 *E. coli* media supplements

LB liquid and LB agar plates for *E. coli* work contains the following antibiotic

Antibiotic	Working concentration	Stock concentration
Ampicillin (+Amp)	100 µg /ml	50 mg /ml
Kanamycin (+Kan)	30 µg/ml	10 mg/ml

#### 2.2.1.4 *Agrobacterium tumefaciens* media supplements

LB liquid and LB agar plates for *A. tumefaciens* contained the following antibiotics

Antibiotic	Working concentration	Stock concentration
Rifampicin (+Rif)	30 µg/ml	4 mg/ml
Kanamycin (+Kan)	30 µg/ml	10 mg/ml

## **2.2.2 Plant culture media**

### **2.2.2.1 1/2 Murashige and Skoog (MS) solid media**

2.2 g of MS mineral salts were combined with water and stirred until completely dissolved. The pH was adjusted to 5.7 with KOH and the final volume was made up to 1 liter with sterile distilled water after 7 g of Agar was added to it. The media was sterilized by autoclaving

### **2.2.2.2 MS media supplements**

MS medium plates were supplemented with the following antibiotic

<b>Antibiotic</b>	<b>Working concentration</b>	<b>Stock concentration</b>
Kanamycin (+Kan)	30 µg/ml	10 mg/ml

## **2.2.3 Antibiotic Stocks**

All antibiotic stocks were sterilized by passing through a 0.22 µm filter, aliquoted into 1.5 ml tubes and stored in -20°C unless otherwise stated

### **2.2.3.1 Ampicillin (50mg/ml)**

1 g of Ampicillin was dissolved in 20 ml of sterile distilled water

### **2.2.3.2 Kanamycin (10mg/ml)**

0.2 g of Kanamycin was dissolved in 20 ml of sterile distilled water

### **2.2.3.3 Rifampicin (4mg/ml)**

0.4 g of Rifampicin was dissolved in 100 ml of methanol and stored in -20°C

## **2.2.4 E. coli Competent Cells preparation buffers**

### **2.2.4.1 SOB**

20 g Tryptone, 5 g Yeast extract, 0.584 g NaCl, 0.186 g KCl, 2.03 g MgCl<sub>2</sub> was dissolved into distilled water completely and made upto 1 liter with distilled water. The buffer was sterilized by autoclaving

#### **2.2.4.2 SOC**

1/100 v/v sterile filtered 2M Glucose was added to 99 ml of sterile autoclaved SOB broth, prior to use

#### **2.2.4.3 TB buffer**

20 ml of 0.5 M stock of Pipes (10 mM), 2.2 g of  $\text{CaCl}_2$  (15 mM) and 18.65 g KCl (250 mM) were mixed into 800 ml of distilled water and the pH adjusted to 6.7 with 1 M KOH. 10.88 g of  $\text{MnCl}_2$  (55 mM) was dissolved into 100 ml of distilled water separately and the solution added to the 800 ml solution from earlier and the final volume was adjusted to 1 liter with distilled water. The solution was sterilized by passing through a 0.22  $\mu\text{m}$  filter and stored at 4°C

#### **2.2.4.4 10 mM Pipes**

7.55 g Pipes was mixed with sterile distilled water, pH was adjusted to 6.4 and made up to 100 ml with distilled water. The buffer was then sterilized by filtration by passing through a 0.22  $\mu\text{m}$  filter

#### **2.2.4.5 TBE Buffer (5X)**

54 g of Tris base (89 mM), 27.5 g Boric acid (89 mM) and 20 ml of 0.5M EDTA (pH 8.0) was completely dissolved in distilled water and made up to 1 liter

#### **2.2.4.6 0.5M EDTA (pH 8.0)**

186.1 g of disodium  $\text{EDTA} \cdot 2\text{H}_2\text{O}$  was dissolved in distilled water and the pH was adjusted to 8.0 using 1M NaOH, before being made up to 1 liter with sterile distilled water

#### **2.2.4.7 TE Buffer (1X)**

1X TE is 10mM Tris HCl and 1mM EDTA. To prepare 1 liter of 1X TE Buffer, 1 ml of 1M Tris HCl (pH 8.0) was combined with 0.2 ml of 0.5M EDTA (pH 8.0) and made up to 1 litre with sterile distilled water

### **2.2.4.8 1M Tris HCl (pH 8.0)**

121.1 g Tris base was dissolved in 700 ml of distilled water and the pH was adjusted to 8.0 by using 3M HCl solution. The total volume was made up to 1 liter with distilled water

## **2.2.5 General reagents preparation**

### **2.2.5.1 X-Gal (50mg/ml)**

0.5 g X-Gal was completely dissolved in 10 ml dimethyl formamide, wrapped in aluminium foil and stored at -20°C

### **2.2.5.2 Isopropylthio-b-D-galactoside (100 mM)**

0.477 g IPTG was dissolved in 20 ml dH<sub>2</sub>O, filter-sterilized through a 0.22 µm filter aliquoted into 1.5 ml tubes and stored at -20°C

### **2.2.5.3 Gel Loading Buffer (6x)**

To make 10 ml of 6x Gel loading buffer, 0.25 mg of Bromophenol blue, 0.25 mg of Xylene cyanol FF and 1.5 g of Ficoll Type 4000 were dissolved in 2.4 ml of 0.5 mM EDTA and made up to 10 ml with sterile distilled water

## **2.2.6 Colored Stains**

### **2.2.6.1 Alexander Stain**

To make Alexander stain the following chemicals/solvents were mixed and made up to 50 ml with sterile distilled water

Ethanol 95%	10 ml
Malachite green (1% in 95% Ethanol)	1 ml
Fuchsin acid (1% in water)	5 ml
Orange G (1% in water)	0.5 ml
Phenol	5 g
Chloral hydrate	5 g
Glacial acetic acid	2 ml
Glycerol	25 ml

### **2.2.6.2 1% Aceto-Orcein**

1 g of orcein was weighed and added to 45 ml of glacial acetic acid. The mixture was boiled until the orcein was dissolved completely. The solution was then slowly added into 55 ml of distilled water and allowed to cool

## **2.2.7 DNA methods**

### **2.2.7.1 Plant genomic DNA extraction**

Genomic DNA was extracted from Wheat tissues using the GenElute™ Plant Genomic DNA Miniprep kit (Sigma G2N70 or G2N350). A small amount of tissue (up to 100 mg) was ground into fine powder in liquid nitrogen. 350 µl of Lysis Solution (Part A) and 50 µl of Lysis Solution (Part B) were added and mixed thoroughly by vortexing and inverting. The mixture was incubated at 65°C for 10 minutes with occasional inversion to dissolve the precipitate. The incubated mixture was removed after 10 minutes and 130 µl of Precipitation Solution was added and mixed thoroughly by inversion for 6-8 times. The sample was then placed on ice for 5 minutes before centrifugation at 13,000 rpm for 5 minutes to pellet the cellular debris, proteins and polysaccharides. The supernatant was carefully pipetted into a GenElute filtration column (blue inset with a 2 ml collection tube) and centrifuged at 13,000 rpm for 1 minute. The filtration column was discarded but the collection tube was retained. To this 700 µl of Binding solution was added directly and mixed thoroughly by inverting the tube 6-8 times. To each GenElute Miniprep Binding Column (with a red O-ring) 500 µl of the Column Preparation Solution was added and centrifuged at 13,000 rpm for 1 minute. The flow-through liquid was discarded after centrifugation. 700 µl of the mixture containing Binding Solution was pipetted onto the Binding column and the tube was centrifuged at 13,000 rpm for 1 minute. The flow-through liquid was discarded after centrifugation and the column was returned to the collection tube. The remaining mixture containing Binding Solution was pipetted onto the Binding column and the tube was centrifuged at 13,000 rpm for 1 minute. The flow-through liquid along with collection tube was discarded after centrifugation. The binding column was placed into a fresh 2 ml collection

tube and to it 500  $\mu$ l of the diluted Wash Solution was added and centrifuged at 13,000 rpm for 1 minute. The flow-through was discarded but collection tube was retained. Another 500  $\mu$ l of Wash Solution was added and the tube was centrifuged at 13,000 rpm for 3 minutes to dry the column. The binding column was then transferred to a fresh 2 ml collection tube and 50  $\mu$ l of pre-warmed (65°C) Elution Solution was applied to it. The tube was centrifuged at 13,000 rpm for 1 minute. The elution step was repeated again with another 50  $\mu$ l of Elution Solution.

When the number of samples from which DNA extractions were made was limited the SIGMA<sup>®</sup> Gen Elute Plant DNA Miniprep kit was used and where the sample numbers were high the DNA was extracted using the DNeasy 96 Plant kit (QIAGEN). Leaves were harvested from 96-well seed trays in the glasshouse and up to 50 mg of each leaf were then placed into each tube in the collection microtube racks provided. The provided plate register cards were used to record the position of each sample in the rack. For best results the leaf material was harvested 24 hours prior to processing and stored at 4°C. One tungsten carbide bead (ball bearing) was added to each collection microtube. 400  $\mu$ l of Buffer AP1 which was preheated to 65°C was combined with 1  $\mu$ l of RNaseA and 1  $\mu$ l of Reagent DX. 400  $\mu$ l of the mixture was pipetted in each collection microtube. Each collection microtube rack was sandwiched between adapter plates and fitted into the tissue lyser (Qiagen) and the samples were ground for 1.5 min at 30Hz. The collection microtube racks were removed and disassembled the collection microtube were centrifuged (Eppendorff, 5810R) for 4 sec or until the speed of the centrifuge reached 3000 rpm. The collection microtube was removed, the caps were discarded and 130  $\mu$ l of buffer AP2 was added to each collection microtube. The microtube racks were sealed with new caps and the racks were shaken vigorously by inverting up and down for 15 sec. The collection microtube racks were centrifuged briefly for 4 sec or until the speed of the centrifuge reached up to 3000 rpm and the racks were incubated at -20°C for 10 min. The collection microtube racks were then centrifuged for 5 minutes at 6000 rpm and the caps were removed and discarded. The supernatant was carefully transferred to new collection

microtube racks in the correct orientation and 600  $\mu$ l of buffer AP3/E added to each sample. The microtube racks were closed with new caps and the racks were shaken vigorously by inverting up and down for 15 sec. The collection microtube racks were centrifuged briefly for 4 sec or until the speed of the centrifuge reached 3000 rpm. Two DNeasy 96 plates were placed on the top of S-Blocks (included in the kit) and the DNeasy 96 plates were marked for identification. 1 ml of each sample from the collection microtube racks were transferred to the DNeasy 96 plates and the plates were sealed with the provided Airpore Tape sheet. The plates were centrifuged for 4 minutes at 6,000 rpm, the tape was removed and 800  $\mu$ l of Buffer AW was carefully added to the DNeasy 96 plates. The plates were centrifuged for 15 minutes at 6,000 rpm to dry the DNeasy membranes. To elute DNA, the DNeasy plates were placed on a provided new rack of Elution microtube RS in the correct orientation and 100  $\mu$ l of Elution buffer AE was added to each collection microtube. The DNeasy plates were sealed with new Airpore Tape sheet and incubated at room temperature for 1 minute. The plate was then centrifuged for 2 minutes at 6,000 rpm to collect the run through and the process was repeated with a further 100  $\mu$ l of buffer AE.

*Arabidopsis* genomic DNA was extracted from young leaves using a Quick DNA extraction protocol. A small amount of leaf tissue (preferably 1-2 rosette leaves) was ground in 400  $\mu$ l of DNA extraction Buffer (200 mM Tris-Cl (pH 7.5), 250 mM NaCl, 25 mM EDTA, 0.5% SDS) (Edwards *et al.*, 1991) with a micropestle in a microcentrifuge tube. The sample was then centrifuged at 14,000 rpm for 5 minutes. 300  $\mu$ l of the supernatant was removed and transferred to a fresh microcentrifuge tube. To this 300  $\mu$ l of isopropanol was added, mixed thoroughly by inverting 3-4 times and the tube was centrifuged at 14,000 rpm for 7 minutes. The supernatant was then completely discarded and the pellet was rinsed with 300  $\mu$ l of 70 % ethanol and was allowed to dry for few minutes. The pellet containing genomic DNA was then resuspended in 30  $\mu$ l of TE buffer (10 mM Tris-HCl (pH 8) and 1mM (EDTA).

### **2.2.7.1.1 Quantification of genomic DNA**

The quantification of genomic DNA was done by gel analysis. The quantification of genomic DNA by agarose gel analysis involved running a known quantity of extracted genomic DNA sample (5  $\mu$ l) on a 1% agarose gel (made with 0.5xTBE) and comparison to a known quantity of uncut Lambda bacteriophage DNA (NEB) ;50ng/ $\mu$ l). As the gel contains ethidium bromide (2ul/100ml; stock 10 mg/ml), the level of fluorescence observed in the test samples was compared with a series of three loadings of the known Lambda sample (10  $\mu$ l, 5  $\mu$ l, 2.5  $\mu$ l). From this, estimates of the concentration of the genomic DNA preps were made.

### **2.2.7.2 Plasmid DNA extraction**

#### **2.2.7.2.1 Plasmid miniprep extraction**

Plasmid DNA from transformed colonies was extracted by using the Genelute Plasmid miniprep DNA kit (sigma PLN70) when large yields were not required (for restriction digests and for use in PCR amplifications). 3 ml of an overnight recombinant *E. coli* culture was centrifuged at 13,000rpm for 2 minutes and the supernatant was discarded. The bacterial pellet was completely resuspended in 200  $\mu$ l of the Resuspension Solution by vortexing and inverting thoroughly until homogenous. The resuspended cells were lysed by adding 200  $\mu$ l of the Lysis Solution and immediately the contents were mixed by gentle inversion (6-8 times) until the mixture becomes clear and viscous. To this, 350  $\mu$ l of Neutralization Solution was added and the tube inverted gently for 4-6 times. The tube was then centrifuged at 13,000 rpm for 10 minutes to precipitate the cell debris. A GenElute Miniprep Binding Column was inserted into a microcentrifuge tube and 500  $\mu$ l of Column Preparation Solution was added to each miniprep column and centrifuged at 13,000 rpm for 1 minute. The flow-through was discarded. The cleared lysate from the neutralization step was then transferred on to the binding column and the tube was centrifuged at 13,000rpm for 1 minute. The flow-through was discarded. 750  $\mu$ l of diluted Wash Solution was added to it and centrifuged at 13,000 rpm for 1 minute to remove residual salt and other contaminants. The flow-through was



discarded and the column was centrifuged for a further 1 minute without adding any further Wash Solution to remove excess ethanol. The column was then transferred to a fresh collection tube and 100  $\mu$ l of nuclease free water was added to the column. The tube was then centrifuged at 13,000 rpm for 1 minute to elute the plasmid DNA.

#### **2.2.7.2.2 Plasmid Midiprep extraction**

The PureYield™ Plasmid Midiprep System (Promega) was used for obtaining large quantities of high quality plasmid DNA (for cloning and transformation purposes). 50 ml of an overnight recombinant bacterial culture was centrifuged at 3700 rpm for 10 minutes to pellet the cells and the supernatant was discarded. The tubes were then dried on a paper towel to remove excess media. The bacterial pellet was completely resuspended in 3 ml of the Cell Resuspension Solution by vortexing and inverting thoroughly until homogenous. The resuspended cells were lysed by adding 3 ml of the Cell Lysis Solution and the contents were mixed by gentle inversion (3-5 times) and incubated at room temperature for 3 minutes. To this 5 ml of Neutralization Solution was added and inverted gently 3-5 times and the tube was incubated for 2-3 minutes at room temperature to allow a white flocculent precipitate to form. A PureYield™ Clearing Column (blue) was placed into a new 50 ml disposable plastic tube and the lysate was transferred into it. The mixture was incubated for 2 minutes and centrifuged at 1500 rpm for 5 minutes. A PureYield™ Binding Column (white) was placed into a new 50 ml disposable plastic tube and the filtered lysate was poured into it. This is centrifuged at 1500 rpm for 3 minutes. 5 ml of Endotoxin Removal Wash solution (with isopropanol added) was added to the PureYield™ Binding Column and the column centrifuged at 1,500 rpm for 3 minutes. The flow-through was discarded and 20 ml of Column Wash Solution (with ethanol added) was added to the PureYield™ Binding Column and centrifuged at 1,500 rpm for 5 minutes. The flow-through was discarded and centrifuged again for an additional 10 minutes to remove any excess ethanol present. Finally the column was then transferred to a fresh 50 ml disposable tube and 600  $\mu$ l of nuclease free water was added into the

PureYield™ Binding column. The tube was then centrifuged at 1500 rpm for 5 minutes to elute the plasmid DNA.

#### **2.2.7.3 Isolation of DNA fragments from agarose gels**

The GelElute™ Gel extraction kit (Sigma) was used to isolate DNA fragments from Agarose gels. Following fractionation and staining of agarose gels using ethidium bromide the desired bands were excised, viewing the gel on a long wave U.V transilluminator (340nm) with a sharp, clean scalpel. The excess gel was trimmed away to minimize Agarose. The gel slice was weighed in a 2 ml Eppendorff tube and 3 gel volumes of Gel Solubilization Solution was added to the gel slice. The gel mixture was incubated at 65°C for 10 minutes, vortexing for every 2-3 minutes during incubation to dissolve the gel. Each of the GenElute Binding Columns needed (G) was inserted into a 2 ml collection tube and 500 µl of the Column Preparation Solution was added and the tube centrifuged at 13,000 rpm for 1 minute. The flow-through liquid was discarded after centrifugation. After the gel slice is completely dissolved 1 gel volume of 100% isopropanol was added and mixed thoroughly by inverting 4-6 times. The sample was then loaded into a GenElute Binding Column in 700 µl portions and centrifuged at 13,000 rpm for 1 minute. The flow-through was then discarded. 700 µl of diluted Wash solution was added to the column and centrifuged at 13,000 rpm for 1 minute. The flow-through was then discarded and centrifugation was repeated again at 13,000 rpm for 1 minute without further addition of Wash solution. The binding column was then transferred to a fresh 2 ml collection tube and 50 µl of pre-warmed Nuclease-free water at 65°C was added to the column and incubated for 1 minute at room temperature. The tube was then centrifuged at 13,000 rpm for 1 minute to elute the DNA.

#### **2.2.7.4 Purification of PCR products**

The PCR products were purified by using QIAquick PCR Purification kit (QIAGEN) for cloning or sequencing purposes. 5 volumes of buffer PB was

added to 1 volume of PCR product and mixed thoroughly by inverting 4-6 times. A QIAquick spin column was inserted into a 2 ml collection tube and the mixture was applied to it. The tube was then centrifuged at 13,000 rpm for 1 minute and the flow-through was discarded. The column was placed into the same tube and 750  $\mu$ l of Wash solution was added to it and the tube centrifuged at 13,000 rpm for 1 minute. The flow-through was then discarded and centrifugation was repeated again at 13,000 rpm for 1 minute without further addition of Wash solution. The QIAquick spin column was then transferred to a fresh 2 ml collection tube and 50  $\mu$ l of Nuclease-free water was added to the column and incubated at 1 minute. The tube was then centrifuged at 13,000 rpm for 1 minute to elute the DNA.

### 2.2.7.5 DNA sequencing

The gel eluted DNA or purified PCR products and or plasmid DNA was sent to Geneservice, Nottingham, UK ([www.geneservice.co.uk](http://www.geneservice.co.uk)) for sequencing using either stock primers at Generservice or gene specific primers. The template DNA for sequencing was prepared according to instructions of Geneservice. The flanking M13F and M13R primers were used to sequence the genes inserted in pGEM T-Easy vector. 35S-seq-F, GFP5-seq-R and T7R primers were used to sequence the gene inserted into pGVT17 vector.

**Table 2.1:** List of primers used for sequencing purposes

Stock primers	Sequence	Length (bp)
M13F	TGTAAAACGACGGCCAGT	18
M13R	CAGGAAACAGCTATGACC	18
35S-seq-F	ATCCCACTATCCTTCGCAAGAC	22
GFP5-seq-R	ATCACCTTCACCCTCTCCACTG	22
T7R	GCTAGTTATTGCTCAGCGG	19

## **2.2.8 RNA methods**

### **2.2.8.1 Total RNA extraction**

Total RNA was extracted from wheat tissues using the Trizol Reagent (Gibco BRL.). A small amount of tissue (100 mg) was ground in 250  $\mu$ l Trizol. After grinding, the volume was made up to 1 ml with further 750  $\mu$ l Trizol and incubated at room temperature for 5 minutes and centrifuged at 11,400 rpm at 4°C for 10 minutes. The supernatant was transferred into a new tube. 200  $\mu$ l of chloroform was added to it, and it was shaken vigorously by hand for 15 seconds before incubation at room temperature for 3 minutes. After centrifuging at 11,400 rpm at 4°C for 15 minutes the aqueous phase was transferred into a new tube. 500  $\mu$ l isopropanol was added, mixed thoroughly and the tube centrifuged again at 11,400 rpm at 4°C for 10 minutes. The supernatant was removed and discarded. 1 ml 75% ethanol was used to wash the pellet by inverting and spinning at 9,000 rpm at 4°C for 5 minutes. The supernatant was removed with a pipette and the pellet was briefly air dried for 5-10 minutes before being resuspended in sterile Milli Q water (10-20  $\mu$ l).

RNA extraction from Arabidopsis tissues was carried out using the RNeasy mini kit (QIAGEN). A small amount of tissue (up to 100 mg) was ground into fine powder in liquid nitrogen. To this powder 450  $\mu$ l of buffer RLT was added and this was mixed thoroughly by vortexing. The lysate was then transferred to a QIAshredder spin column (lilac) placed in a 2 ml collection tube, and centrifuged for 2 min at 13,000 rpm. The supernatant was then carefully pipetted into a new 2 ml collection tube without disturbing the pellet. To this 225  $\mu$ l of Absolute ethanol was added and mixed thoroughly by pipetting. The sample (usually 650  $\mu$ l), including any precipitate that may have formed was transferred to an RNeasy spin column (pink) placed in a 2 ml collection tube and centrifuged for 15 s at 10,000 rpm. The flow-through was discarded. To this 700  $\mu$ l of Buffer RW1 was added to wash the column and the tube was centrifuged for 15 s at 10,000 rpm and the flow-through discarded. 500  $\mu$ l Buffer RPE was added to the RNeasy spin column and centrifuged for 15 s at 10,000 rpm to wash the spin column

membrane and the flow-through discarded. Another 500 µl Buffer RPE was added to the RNeasy spin column and centrifuged for 2 min at 10,000 rpm to further wash the spin column membrane. The RNeasy spin column was inserted into a new 1.5 ml collection tube and 50µl RNase-free water was added directly to the spin column membrane and centrifuged for 1 min at 10,000 rpm to elute the RNA.

#### **2.2.8.2 Quantification of RNA**

The concentrations of RNA and plasmid DNA samples were determined using the Nanodrop machine. Around 1.5 µl of sample was loaded on the nanodrop machine after blanking the spectrophotometer with distilled water or other appropriate buffer and readings were noted.

#### **2.2.9 First strand cDNA synthesis**

##### **2.2.9.1 DNase treatment of RNA**

RNA was subjected to DNaseI treatment to digest any contaminating DNA. 20 µl of RNA was mixed with 2.5 µl 10x Turbo DNase Buffer, 0.5 µl Turbo DNase and 2 µl nuclease free distilled water. The above mixture was incubated at 37°C for 30 minutes. The enzyme was inactivated by the addition of 5 µl of DNase inactivation reagent and incubated at room temperature for 2 minutes. The mixture was then centrifuged at 10,000 rpm at room temperature for 1 minute and the supernatant transferred to a new tube.

##### **2.2.9.2 Reverse Transcription**

The DNase treated RNA was used for first strand cDNA synthesis using Superscript III Reverse Transcriptase (Invitrogen). 10 µl of DNase treated RNA was mixed with 1.0 µl of 50mM Oligo dT<sub>18</sub> (Invitrogen), 1.0 µl of 10mM dNTP mix (NEB) and 8 µl of RNase free water and the mixture was heated for 5 min at 65°C in a PCR machine. The mixture was then collected by centrifugation and 4 µl of 5x first strand Buffer, 1 µl of 0.1M dTT, 1 µl of RNase out and 1 µl Superscript III Reverse Transcriptase enzyme

(Invitrogen) was added to it. The mixture was then incubated at 50°C for 1 hour and inactivated at 70°C for 15 minutes in a PCR machine.

## **2.2.10 PCR (Polymerase Chain Reaction) methods**

### **2.2.10.1 Gradient PCR method**

Gradient annealing temperature optimization PCR was carried out by using Standard Taq DNA polymerase (NEB). Each individual reaction mixture consists of 200  $\mu$ M of each dNTP, 0.25  $\mu$ M of each primer, 10-50 ng of template DNA, 1 x Standard Taq Buffer (including  $\text{MgCl}_2$ ; final 1.5 mM  $\text{MgCl}_2$ ) and 1 U of Taq DNA polymerase in a 20  $\mu$ l volume. Each PCR reaction was performed in a single PCR tube. Reactions were carried out using the Thermo Hybaid gradient PCR with an annealing temperature gradient ranging from 50-65°C and the rest of the conditions were the same as in Table 2.2 for Standard Taq DNA polymerase.

### **2.2.10.2 General and Phusion PCR**

Amplification of DNA was performed using either Standard Taq DNA polymerase (NEB) or Phusion polymerase (Finnzymes). Standard Taq DNA polymerase was used for routine amplifications and Phusion polymerase was used where higher fidelity is needed (for cloning and sequencing). For standard Taq DNA polymerase, 200  $\mu$ M of each dNTP, 0.25  $\mu$ M of each primer, 10-50 ng of template DNA, 1 x Standard Taq Buffer (including  $\text{MgCl}_2$ ; 1.5 mM  $\text{MgCl}_2$ ) and 1 U of Taq DNA polymerase in a 20  $\mu$ l volume for each reaction. For Phusion, 200  $\mu$ M of each dNTP, 0.5  $\mu$ M of each primer, 10-50 ng of template DNA, 1 x HF Buffer (including 7.5 mM  $\text{MgCl}_2$  final concentration) and 1 U of Phusion polymerase in a 50  $\mu$ l volume. The PCR reaction conditions for Standard Taq DNA polymerase and Phusion polymerase are as shown in Table 2.2.

**Table 2.2** Reaction conditions for Standard and Phusion PCR

PCR step	Standard Taq Polymerase	Phusion polymerase
Initial denaturation	94°C for 3 min	98°C for 30 sec
Denaturation	94°C for 30 sec	98°C for 10 sec
Annealing	T <sub>m</sub> °C for 30 sec	T <sub>m</sub> °C for 30 sec
Extension	72°C for 30sec/0.5kb	72°C for 30sec/0.5kb
Final Extension	72°C for 7 min	72°C for 7 min
No of cycles	35	30

(T<sub>m</sub>°C – depends on the annealing temperature of the primer pair)

Optimum annealing temperature of the primer pair was determined by performing a Gradient PCR.

### 2.2.10.3 Long Amp Taq DNA PCR

Long Amp PCR system (NEB) was used for the isolation of the full length cDNA and genomic sequence of *TaRAD51* and *TaDMC1* genes using Highbury gDNA. Each PCR reaction consists of 200 µM of each dNTP, 0.4 µM of each primer, 10-50 ng of template cDNA, 1 x Long Amp Taq reaction Buffer and 1 U of Long Amp Taq DNA polymerase in a 25 µl volume. The PCR cycle conditions used were; denaturation at 94°C for 30 sec, then 35 cycles of 94°C for 10 sec, T<sub>m</sub>°C (depends on the annealing temperature of the primer pair) for 1 min, 65°C for 30sec to X min (depending on the size of the fragment; for every 500bp a 30sec extension time was used) and a final elongation step of 65°C for 10 min.

### 2.2.10.4 Colony PCR

Selection of the transformed *E. coli* and *Agrobacterium* colonies with plasmid vectors containing the desired insert was carried out by colony PCR using gene specific primers. The general PCR conditions for an individual reaction were: 0.1 µg plasmid DNA (or 1 µl bacterial culture), 1.5 mM MgCl<sub>2</sub> (NEB); 200 µM of each dNTP (NEB); 0.25 µM of each primers; 1 U of Taq DNA polymerase (NEB) in 20 µl of 1X Reaction Buffer (NEB).

PCR reactions were carried out in a Thermal cycler using the stated conditions in Table 2.2

#### **2.2.10.5 RT-PCR**

Gene expression analysis was performed by RT-PCR at 35 PCR cycles. Each PCR reaction consists of 200  $\mu\text{M}$  of each dNTP, 0.5  $\mu\text{M}$  of each primer, 10-50 ng of template cDNA, 1 x HF Buffer (including 7.5mM  $\text{MgCl}_2$ ) and 1 U of Phusion polymerase in a 50  $\mu\text{l}$  volume. The PCR cycle conditions used were; denaturation at 94°C for 3min, then 30 cycles of 94°C for 30s,  $T_m$ °C (depends on the annealing temperature of the primer pair) for 30s, 72°C for 30s to 1 min (depending on the size of the fragment, for every 500bp a 30 sec extension time was used) and a final elongation step of 72°C for 7 min.

#### **2.2.10.6 Quantitative real-time PCR (Q-PCR)**

Q-PCR amplification reactions were performed in the LightCycler® 480 instrument (Roche Applied Science, UK) using the 384 block system with Brilliant® SYBR® Green QPCR Master Mix (QIAGEN). Q-PCR reactions were done in triplicates and were set up in a total volume of 12  $\mu\text{l}$  containing 5  $\mu\text{l}$  of cDNA (diluted to 1:10 in SDW), 6  $\mu\text{l}$  of 2X master mix and 1  $\mu\text{l}$  of an equal mix of forward and reverse primers. The reaction conditions for QRT- PCR were as follows: 15 min at 95°C followed by 45 cycles of 30s at 95°C, 30s at 60°C and 30s at 72°C. A melt curve was obtained from the PCR product at the end of amplifications by heating from 70°C to 99°C.

#### **2.2.11 Agarose Gel Electrophoresis**

Agarose gels of appropriate concentration (generally 1%) were cast using agarose melted in 0.5x Tris-borate EDTA (TBE) buffer in a microwave oven. The molten gel was allowed to cool to 50-60°C and gels were supplemented with ethidium bromide at a final concentration of 0.5 $\mu\text{g}/\text{ml}$ . The molten agarose was then poured into a gel tray with its edges taped and an appropriate comb or combs inserted. After the bubbles were removed the



gel tray containing the agarose was then left at room temperature until it had properly set. The gel tray was then submerged in a gel tank containing 0.5X TBE buffer. The wells were loaded with DNA samples after mixing them with an appropriate volume of 6x gel loading dye. After loading appropriate size markers and samples, gels were run at 90 V for 1 hr unless otherwise stated. The gels were then visualized through the U.V. transilluminator and photographed using a video copy processor.

Size standards used were: (All from NEB)

- 1Kb ladder (10Kb-500bp, ~1kb intervals between bands)
- 100bp ladder (1.5Kb – 100bp, 100bp intervals)
- 2-log ladder (10kb-100bp, 100bp then 1kb intervals)

## **2.2.12 DNA cloning methods**

### **2.2.12.1 Restriction digestions for cloning**

The plasmid DNA extracted from overnight cultured bacterial colonies was digested with restriction endonucleases supplied by NEB. The restriction conditions and selection of restriction enzymes differed with the plasmid DNA (vector) to be digested and inserts. The restriction conditions for pGEM T-Easy vector system for individual reactions in a total volume of 10  $\mu$ l include 5  $\mu$ l (0.5  $\mu$ g) plasmid DNA, 1  $\mu$ l (1X) of 10X Buffer 3, 1  $\mu$ l (1X) of 10X BSA and 0.2  $\mu$ l of (2U) of *NotI* or *EcoRI* restriction enzyme. The reaction was incubated for 2hrs at 37°C and run on 1% Agarose gel at 90v for 1hr.

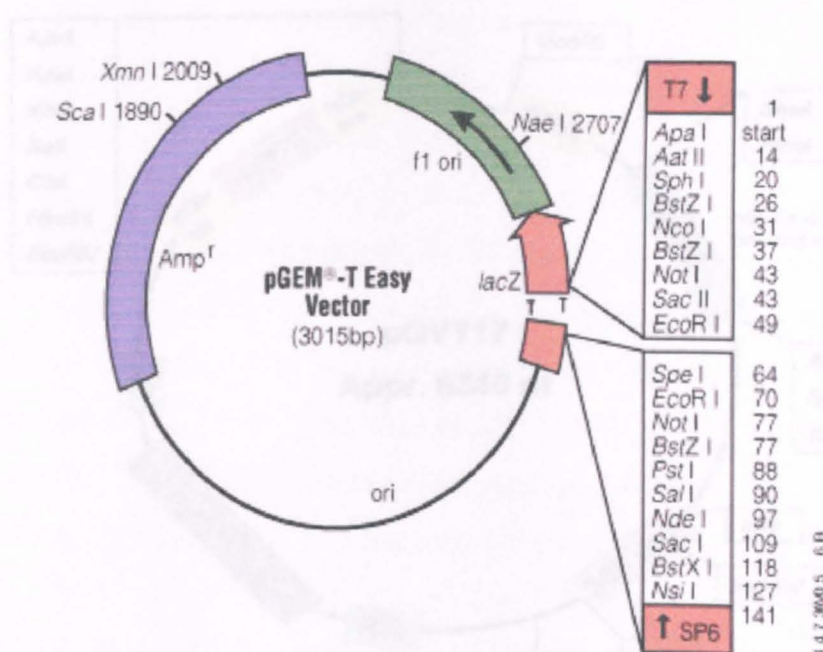
The restriction conditions for the pGVT17 vector system for individual reactions in a total volume of 100  $\mu$ l include approximately 25  $\mu$ l (0.5  $\mu$ g) plasmid DNA, 10  $\mu$ l (1X) of 10X Buffer 4, 10  $\mu$ l (1X) of 10X BSA and 1 U of *XbaI* and *AscI/PmeI* restriction enzymes. The reaction was incubated for 3hr at 37°C and a sample run on 1% Agarose gel at 90v for 1hr.

**Table 2.3** List of restriction enzymes used for cloning into pGEM T-Easy and pGVT17 vectors

Restriction Enzyme	NEB Buffer Used+BSA	Type of cutter	Recognition site
<i>EcoRI</i>	3	Cohesive end	5' ... GAATTC ... 3' 3' ... CTTAAG ... 5'
<i>NotI</i>	3+BSA	Cohesive end	5' ... GCGGCCGC ... 3' 3' ... CGCGGCCG ... 5'
<i>XbaI</i>	4+BSA	Cohesive end	5' ... TCTAGA ... 3' 3' ... AGATCT ... 5'
<i>AscI</i>	4	Cohesive end	5' ... GGCGGCCG ... 3' 3' ... CGCGGCCG ... 5'
<i>PmeI</i>	4	Blunt end	5' ... GTTTAAAC ... 3' 3' ... CAAAATTG ... 5'

### 2.2.12.2 Ligation into the pGEM T-Easy vector system

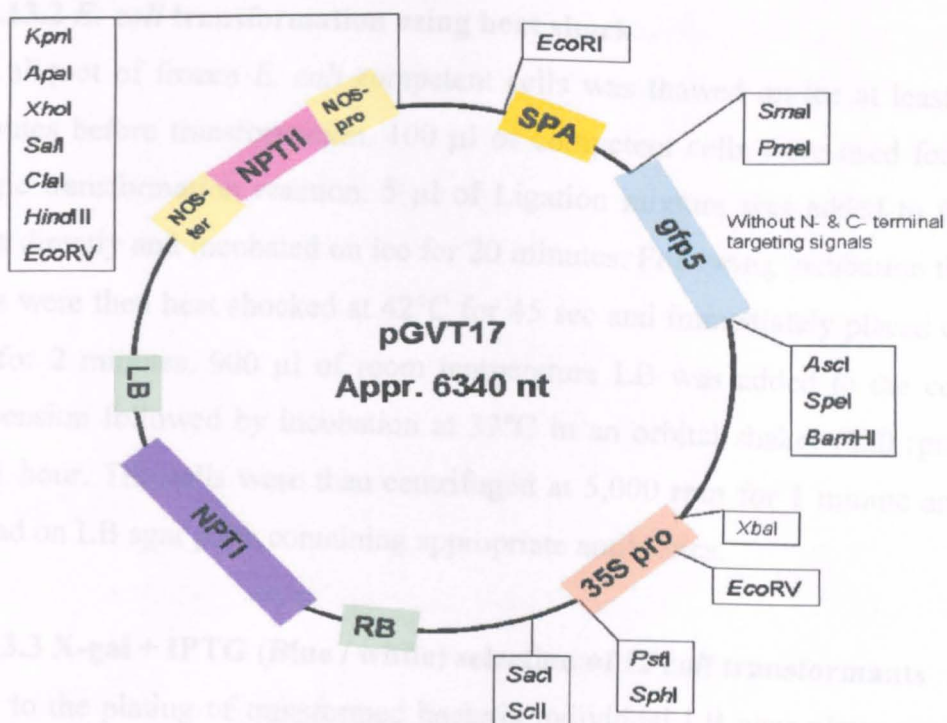
The purified PCR products were ligated into the pGEM T-Easy vector system (Promega, UK) for sub-cloning and sequencing purposes (Figure 2.1). The ligation mixture was composed of a total reaction volume of 10  $\mu$ l containing a molar ratio of ~3:1 insert to vector DNA, 0.5  $\mu$ l of vector (25 ng), 5  $\mu$ l of ligation buffer (1x) and 0.5  $\mu$ l of T4 DNA ligase. The mixture was then incubated at room temperature for 1 hour or kept at 4°C overnight to obtain more transformants after which the ligation mixture were directly used for transformation.



**Figure 2.1** pGEM T-Easy vector restriction map and sequence reference points ([www.promega.com](http://www.promega.com)).

### 2.2.12.3 Ligation into the pGVT17 vector system

The pGVT17 vector system was used for transforming wheat meiotic genes into Arabidopsis (Figure 2.2). The Quick Ligation Kit (NEB, UK) was used for cloning of purified PCR products amplified by the proof-reading polymerase (Phusion polymerase) into the pGVT17 vector system. The ligation mixture was composed of a total reaction volume of 21  $\mu$ l containing 50ng of vector with the appropriate amount of DNA to obtain a molar ratio of ~3:1 insert to vector DNA and made up to 10  $\mu$ l with distilled water, 10  $\mu$ l of 2x Ligation buffer and 1  $\mu$ l of T4 DNA ligase. The mixture was thoroughly mixed and incubated at room temperature for 30 minutes and chilled on ice after which it was directly used for transformation.



**Figure 2.2** Restriction map of the pGVT17 binary vector used for cloning and transformation purposes (Figure courtesy of Dr. Philip Vain, JIC)

## 2.2.13 *E. coli* methods

### 2.2.13.1 *E. coli* competent cells preparation

The *E. coli* competent cells were prepared for plasmid transformation according to the Inoue *et al.*, (1990) protocol. DH5 $\alpha$  cells were streaked out from a glycerol stock and grown on an LB plate, incubated overnight at 37°C. A single colony was picked and inoculated into LB liquid media and grown at 220 rpm in a 37°C rotary shaker incubator for 16-18 hours. Cells were diluted to 1/100 in SOB and grown at 37°C until OD<sub>600</sub> = 0.6. Cells were then pelleted at low speed (2500 g) for 5 min at 4°C. The supernatant was removed and the pellet was resuspended gently in 80 ml TB Buffer and placed on ice for 10 min. Cells were pelleted again as before after adding 1.4 ml DMSO to 18.6 ml ice-cold TB buffer. The supernatant was removed, cells were resuspended in 20 ml TB-DMSO (7% DMSO in TB) and placed on ice for 10 min. Cells were then aliquoted in batches in pre-cooled tubes on ice and flash frozen immediately in liquid nitrogen and stored in -80°C.

### **2.2.13.2 *E. coli* transformation using heat shock**

An aliquot of frozen *E. coli* competent cells was thawed on ice at least 5 minutes before transformation. 100 µl of competent cells were used for a single transformation reaction. 5 µl of Ligation mixture was added to the cells directly and incubated on ice for 20 minutes. Following incubation the cells were then heat shocked at 42°C for 45 sec and immediately placed on ice for 2 minutes. 900 µl of room temperature LB was added to the cell suspension followed by incubation at 37°C in an orbital shaker (220 rpm) for 1 hour. The cells were then centrifuged at 5,000 rpm for 1 minute and spread on LB agar plate containing appropriate antibiotics.

### **2.2.13.3 X-gal + IPTG (Blue / white) selection of *E. coli* transformants**

Prior to the plating of transformed bacteria individual LB agar plates were supplemented with 40 µl of X-gal (20 mg/ml) and 40 µl of 100 mM IPTG which was spread onto the plates and the plates were then incubated upside down in 37°C for 30 minutes to dry. After plating of transformed bacteria onto the plates, they were allowed to dry for 30 minutes, before incubation overnight at 37°C in an inverted position. After storage at 4°C for 2 hours, positive, white colonies were picked up for further analysis by colony PCR.

### **2.2.13.4 Glycerol Stock preparation**

Glycerol stocks were prepared by adding 0.5 ml of the overnight liquid culture after plasmid preparation to an equal volume (0.5 ml) of 30% (v/v) glycerol in LB medium. The tube containing the glycerol stock was then mixed by inversion and snap frozen in liquid nitrogen and stored at -80°C.

## **2.2.14 *Agrobacterium tumefaciens* methods**

### **2.2.14.1 Preparation of *Agrobacterium* electrocompetent cells**

An overnight 5 ml of starter culture (in LB broth) with a single colony of the EHA105 strain of *Agrobacterium* was grown at 220 rpm in 28°C rotary shaker incubator for 16-18 hours. 200 ml of LB with appropriate antibiotics was inoculated with the 1 ml of overnight culture of *Agrobacterium* and the culture was incubated at 28°C with vigorous agitation until the cells have

reached  $OD_{600}=0.8$ . The culture was chilled by standing it in an ice-water bath. The cells were then pelleted in bench top centrifuge at 5000 rpm for 10 minutes at 4°C in a pre-chilled rotor. The supernatant was discarded and 5-10 ml of ice-cold sterile distilled water was added and the pellet was pipetted with a wide-bore pipette gently up and down until no clumps were remained. The volume was adjusted to 50 ml with ice-cold distilled water. The above step was repeated twice before resuspension of the pellet in 5 ml of 1 X TE buffer. The cells were then dispensed into 50 µl aliquots and snap frozen in liquid nitrogen and stored in -80°C.

#### **2.2.14.2 Transformation of *Agrobacterium* by Electroporation**

Transformation of *Agrobacterium* electrocompetent cells was performed according to the recommendations supplied with the Genepulsor (Biorad). 2.5 µl of plasmid DNA was mixed with 50 µl of *Agrobacterium* competent cells that were thawed on ice for at least 5 minutes before transformation and the mixture was transferred to an ice-cold disposable electroporation cuvette (2mm electrode gap) and subject to electroporation with the following conditions : Capacitance: 25µF, Voltage: 2.4V, Resistance: 200 Ohm and Pulse length: 5 msec. Immediately after electroporation, the cells were mixed with 1 ml of LB and the bacterial suspension was transferred to a fresh 1.5 ml eppendorf tube. The cells were then incubated on a rotary shaker at 28°C for 3 hours with gentle agitation. The cells were then centrifuged at 5,000 rpm for 1 minute and spread onto LB agar plate containing the appropriate antibiotic for the T-DNA vector and incubated overnight at 28°C incubator for 2-3 days.

#### **2.2.15 Wheat methods**

##### **2.2.15.1 *In vitro* germination of wheat seeds**

The wheat seeds were first surface sterilized with 5% Parazone for 10 min and rinsed thoroughly with sterile distilled water three times before placing them onto 120 mm<sup>2</sup> Petri dishes containing a 110 mm Whatman No. 1 filter paper and 2 ml sterile distilled water. The dishes were sealed with parafilm

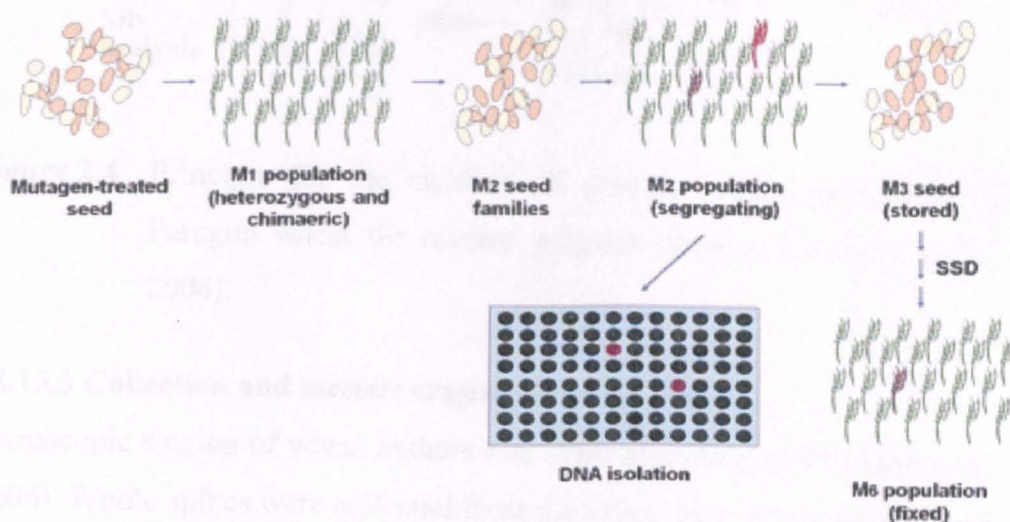


and kept flat in a growth room at 22°C for 3-5 days before transferring them to soil.

## 2.2.15.2 Mutagenesis methods

### 2.2.15.2.1 Fast-neutron mutagenesis

The Highbury wheat deletion lines were generated *de novo* by subjecting 3000 seeds of Highbury (CS5B) to a fast-neutron irradiation dose of 3.0 GyNf at the International Atomic Energy Agency's laboratories (Seibersdorf, Austria; protocol no.FN662). The M<sub>1</sub> generation was grown in a greenhouse, allowed to self-fertilize, and the seed was harvested at John Innes Centre (Norwich, UK). The M<sub>2</sub> and successive generations were produced by self-fertilization under single-seed descent method (Figure 2.3).



**Figure 2.3** General strategies for the production of Highbury wheat deletions lines using fast-neutron irradiation performed at John Innes Centre (Norwich, UK).

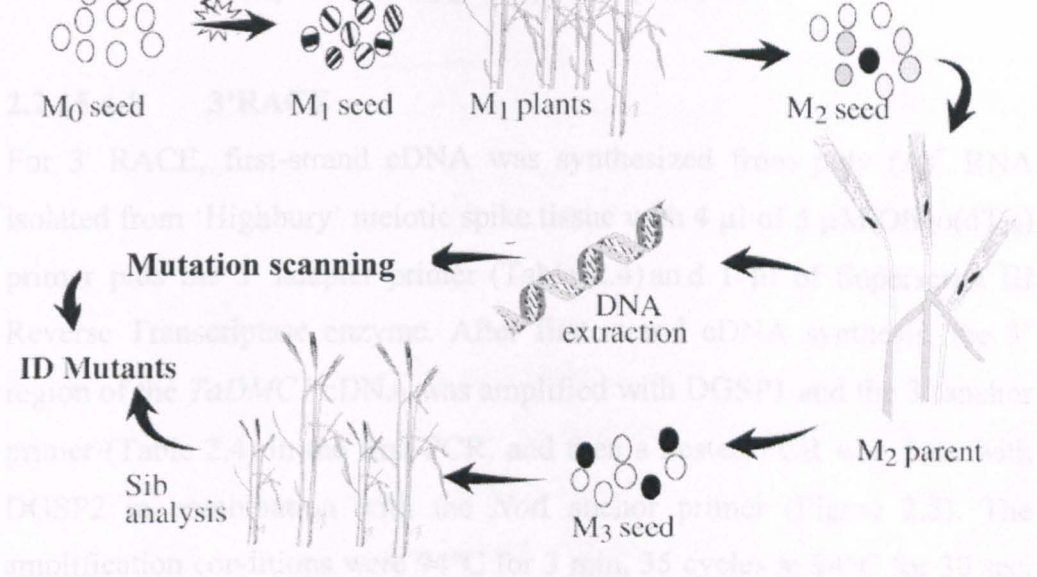
#### 2.2.15.2.2 Gamma ray mutagenesis

Seeds of *Triticum aestivum* cultivar Paragon (RAGT's NABIM Group 1 variety) were used as starting material for the creation of gamma mutant populations. Paragon wheat seeds were mutagenized with 25-250 Grays doses of Gamma rays (Figure 2.4). The mutagenized seeds were grown to M<sub>1</sub> plants at John Innes Centre (Norwich, UK) and because M<sub>1</sub> plants are

heterozygous for mutations, the population was taken to M<sub>2</sub> lines by self-pollination. M<sub>3</sub> seed from individual M<sub>2</sub> plants was harvested and catalogued.

#### 2.2.15.4 Rapid Amplification of cDNA ends (RACE)

RACE was performed for the amplification of the 3' and 5' ends of the TaDMS1 cDNA sequence using a 5' primer.



**Figure 2.4** Principle for the creation of gamma mutant population of Paragon wheat for reverse genetics (Source: Caldwell *et al.*, 2004).

#### 2.2.15.3 Collection and meiotic staging of wheat anthers

Microscopic staging of wheat anthers was done according to Crismani *et al* (2006). Whole spikes were collected from the wheat plant when the flag leaf had emerged by about 2 to 4 cm above the last true leaf. Single spikelets were dissected from the rachis and all the three anthers were removed from the primary floret using syringes. Two anthers were placed on a Petri-dish on ice and the remaining anther was placed on microscopic slide and stained with a 15 µL drop of 1% aceto-orcein for 5 minutes or until a dark ring is formed on the edges of the drop. The anther was then sliced transversely using sharp scalpel and a cover-slip was placed over it. The anther halves were squashed gently but firmly to release the meiocytes. The meiocytes were observed using a compound microscope under 40X magnification. Around 7-10 anthers were pooled according to stages, snap frozen in liquid



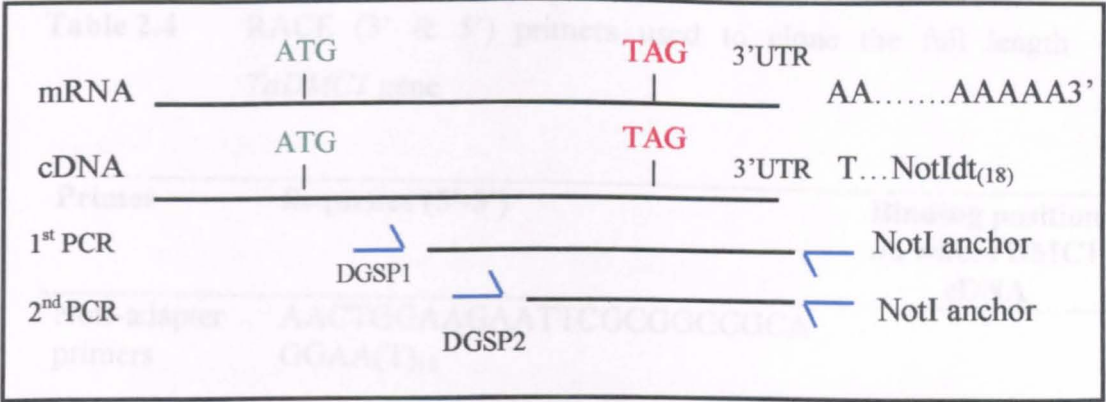
nitrogen and stored at -80°C for later use. For best results spikes were collected early in the morning or mid-morning.

#### **2.2.15.4 Rapid Amplification of cDNA ends (RACE)**

RACE was performed for the amplification of the 3' and 5' ends of the *TaDMC1* cDNA sequence using gene specific primers.

##### **2.2.15.4.1 3'RACE**

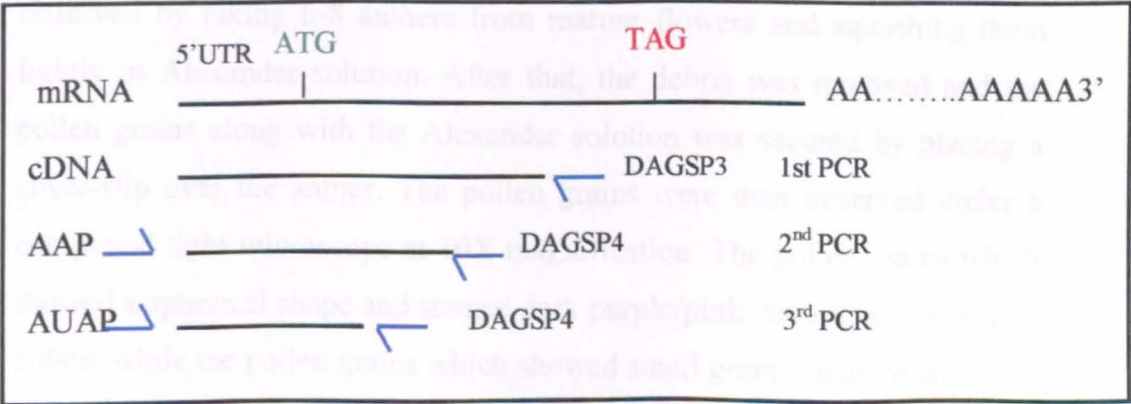
For 3' RACE, first-strand cDNA was synthesized from poly (A)<sup>+</sup> RNA isolated from 'Highbury' meiotic spike tissue with 4 µl of 5 µM Oligo(dT)<sub>18</sub> primer plus the 3' adapter primer (Table 2.4) and 1 µl of Superscript III Reverse Transcriptase enzyme. After first strand cDNA synthesis, the 3' region of the *TaDMC1* cDNA was amplified with DGSP1 and the 3' anchor primer (Table 2.4) in the first PCR, and then a nested PCR was done with DGSP2 in combination with the *NotI* anchor primer (Figure 2.5). The amplification conditions were 94°C for 3 min, 35 cycles at 94°C for 30 sec, 55°C for 30 sec, 72°C for 1 min, with a final extension of 7 min at 72°C and the nested PCR was done with 1:1000 diluted 1<sup>st</sup> PCR as template and the amplification conditions were 94°C for 3 min, 35 cycles at 94°C for 30 sec, 55°C for 30 sec, 72°C for 1 min, with a final extension of 7 min at 72°C. The amplified fragments after the 2<sup>nd</sup> nested PCR were then gel eluted and cloned into the pGEM-T Easy vector system (Section 2.2.10.1) for sequencing.



**Figure 2.5** General scheme for the amplification of 3' ends of *TaDMC1* gene by 3'RACE.

**2.2.15.4.2 5' RACE**

The 5' of the *TaDMC1* transcript was obtained with an Invitrogen 5' RACE kit. The first strand cDNA was synthesized from Highbury meiotic spike tissue with the gene specific primer DGSP3 (Table 2.4) and 1 µl of Superscript III Reverse Transcriptase enzyme (Invitrogen). The specific cDNA was then amplified by two rounds of PCR with the gene-specific nested primers DAGSP4 and DAGSP5 along with primers provided in the kit (Figure 2.6).



**Figure 2.6** General scheme for the amplification of 5' ends of *TaDMC1* gene by 5' RACE.

**Table 2.4** RACE (3' & 5') primers used to clone the full length *TaDMC1* gene

Primer	Sequence (5'-3')	Binding position on wheat DMC1 cDNA
NotI-adaptor primers	AACTGGAAGAATTTCGCGGCCGCA GGAA(T) <sub>18</sub>	-
NotI-anchor primer	AACTGGAAGAATTTCGCGG	-
DGSP1	GAAGAGCCTTTCAGGCTTCTG	700
DGSP2	GGTGAACCTTGCAGAGCGCCAG	772
DGSP3	CTGGCGCTCTGCAAGTTCACC	772
DGSP4	GGAAGACCCAGTTGGCTCAT	461
DGSP5	GGTGAACCTTGCAGAGCGTCAG	207

#### 2.2.15.5 Alexander staining

The evaluation of pollen viability of both wild type and mutant plants was performed using Alexander stain. The viability of pollen grains was observed by taking 6-8 anthers from mature flowers and squashing them lightly in Alexander solution. After that, the debris was removed and the pollen grains along with the Alexander solution was secured by placing a cover-slip over the anther. The pollen grains were then observed under a compound light microscope at 10X magnification. The pollen grains which showed a spherical shape and stained dark purple/pink were scored as viable pollen, while the pollen grains which showed small grains, aborted and were light green were scored as non-viable or sterile pollen. Pollen viability of each line was calculated from the average scores of pollen viability across six plants.

#### **2.2.15.6 Cytogenetic analysis**

Wheat spikes from both wild type and mutants at appropriate stages of meiosis were collected (section 4.2.2) and individual ears were fixed in 80 ml of a mixture containing ethanol and Acetic acid (3:1) for 24 hours at room temperature. The wheat spikes were then washed three times with 70% ethanol and stored in 70% ethanol at 4°C for future use. Meiosis in anthers of pollen mother cells were observed by placing three to four anthers onto a microscopic slide and squashing them lightly in 1% Aceto-orcein. The squashed anther debris was removed and the microspores were secured by placing a cover-slip on top. The microspores were then observed under the compound light microscope at 40X magnification.

#### **2.2.15.7 Fertility assessment**

Wheat ears were harvested from the plants growing in the field and the glasshouse and phenotypic characterization of both wild-type and mutant plants were performed. Recorded were; number of spikelets/ear, number of grains/ear, number of grains/spikelet and average grain weight. Ear fertility was calculated as follows: average of the number of grains/spikelet in six ears. To estimate the number of grains/spikelet, the number of grains/ear was divided by spikelets/ear.

### **2.2.16 Arabidopsis methods**

#### **2.2.16.1 *In vitro* germination of seeds**

To grow plants under *in vitro* conditions, Arabidopsis seeds were placed in a 1.5 ml microcentrifuge tube and sterilized with 10% (V/V) bleach, and mixed by inversion. Seeds were then washed (3 times) with sterile water containing 0.01% (V/V) Triton X-100. Finally, water-diluted ethanol at 70% (V/V) was added to the tubes and all seeds were poured onto filter papers, and allowed to dry. Seeds were spread onto ½ MS (0.8% agar, pH 5.8) (with or without antibiotic selection). The plates were placed at 4°C for 2 days for

stratification and to synchronize germination and transferred to standard growing conditions thereafter.

#### **2.2.16.2 Transformation of *Arabidopsis* by floral dipping**

The transformation of *Arabidopsis* was done according to the floral dipping method of Clough and Bent (1998) using *Agrobacterium* strain EHA105. The plants were grown for 3 weeks under the above conditions before actual transformation. After 3 weeks the emerging bolts were clipped to induce the growth of secondary bolts. Two days prior to plant transformation the *Agrobacterium* strain carrying appropriate binary vector was grown in 5 ml of LB with appropriate antibiotics for 16-18 hours at 28°C. 2-3 ml of the overnight grown culture was then inoculated into 400 ml of LB with appropriate antibiotics and grown until OD<sub>600</sub>=0.8 or higher. The bacterial suspension was then spun down and resuspended in 5% Sucrose (freshly prepared). Silwett L-77 was then added to the resuspended culture at a final concentration of 0.05% and the *agrobacterium* suspension was transferred to a 1000 ml plastic beaker for dipping plants. The aerial parts of flowering *Arabidopsis* were then dipped into *Agrobacterium* solution for 20 sec with gentle agitation. The same suspension was used to dip 10-12 plants. After dipping, the plants were then covered with plastic sleeves for 24 hours to maintain high humidity. After 24 hours the plastic sleeves were opened and the plants were watered and grown normally. The dried siliques from the transformed plants were then harvested after 2-3 weeks.

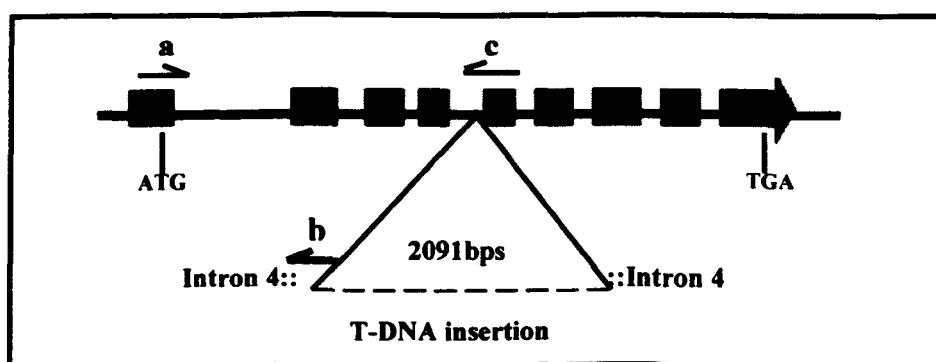
##### **2.2.16.2.1 Screening of *Arabidopsis* transformants**

The dried siliques from the transformed plants were harvested and surface sterilized (section 2.2.16.1). The seeds were then plated on ½ MS media containing 30 µg/ml of Kanamycin. The Petri dishes were then cold treated for 2 days in the dark to break dormancy and then transferred to the growth room with 16/8hr photoperiod at 23°C. After 2 weeks, Kanamycin resistant plants with true leaves and extended root system were then transferred into 9 cm pots.

### 2.2.16.3 Genotypic analysis of the *Arabidopsis rad51* and *dmcl* T-DNA mutants

#### 2.2.16.3.1 Genotypic analysis of the *Arabidopsis rad51* T-DNA lines

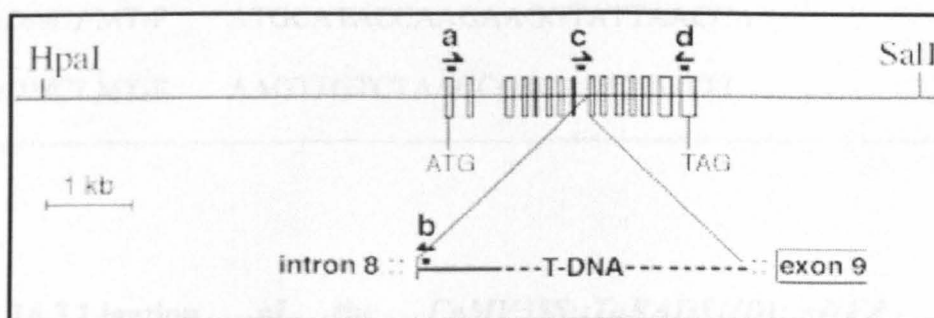
Genotyping of the *Arabidopsis rad51* T-DNA lines was done using a PCR-based approach using primer information provided by Dr. Marie-Pascale Doutriaux (Paris-Sud University, France). Genomic DNA was isolated (section 2.2.7) from each plant within the line and used as the template for PCR amplification corresponding to wild-type allele and the insertion allele. For the wild-type allele amplification, primer 'a' and primer 'c' (Figure 2.7) were used. These are designed based on the *AtRAD51* genomic sequence in the database (NM\_122092). Primer 'a' was used with primer 'b' (which was designed from the LB region of the T-DNA) to confirm the presence of the T-DNA (Figure 2.7). PCR was performed separately with the two different sets of primers (a & c; a & b) (Table 2.5) on single plant *Arabidopsis* genomic DNA and the PCR products were mixed and run on an agarose gel to identify the genotype of the *Arabidopsis* plant.



**Figure 2.7** Schematic representation of the *AtRAD51* T-DNA insertion site. Exons are indicated by closed boxes. Primers a, b and c used to define plant genotypes.

### 2.2.16.3.2 Genotypic analysis of the *Arabidopsis dmc1* T-DNA lines

To genotype the *Arabidopsis* lines carrying the T-DNA insertion in the *DMC1* gene, PCR was done using primer information provided from Dr. Marie-Pascale Doutriaux, as indicated below. The DNA was prepared from the leaves of mature *Arabidopsis* plants growing in soil and amplified using primers derived from the genomic sequence of *AtDMC1*(NM\_113188): the 'a' primer and the 'c' primer were used to confirm the presence of the wild-type allele. The 'b' primer was designed to the LB region of the T-DNA and used in conjunction with 'a' primer to confirm the presence of the T-DNA insertion allele (Figure 2.8). PCR was done separately with the two different sets of primers (a & c; a & b) (Table 2.5) on *Arabidopsis* genomic DNA and the PCR products were mixed and run on an agarose gel to identify different genotypes of *Arabidopsis*.



**Figure 2.8** Schematic representation of the *AtDMC1* genomic region. Exons are indicated by open boxes. Primers a, b and c used to define plant genotypes as described in text.

**Table 2.5** Primers used for the genotypic characterization of the *AtDMC1* and *AtRAD51* T-DNA insertion lines

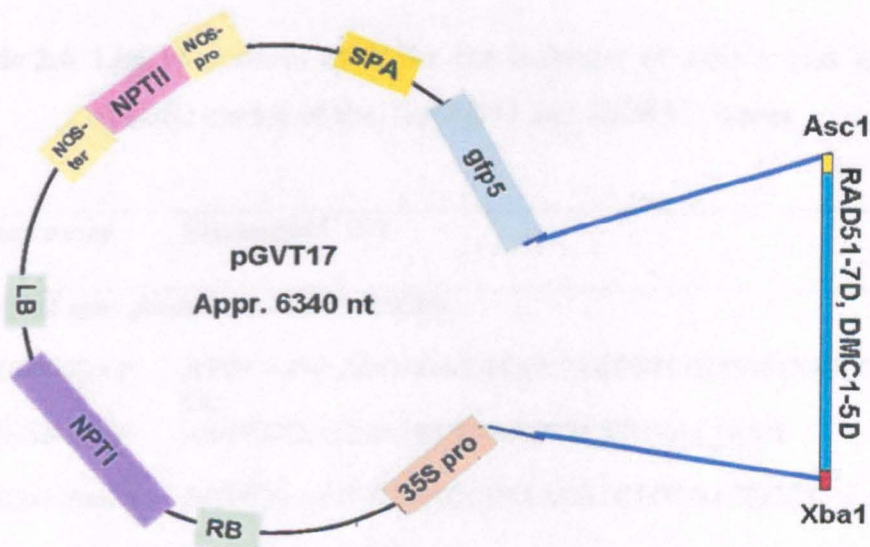
Primer name	Sequence(5'-3')	Tm (°C)
<b><i>AtRAD51</i> WT and T-DNA primers</b>		
<i>AtRAD51</i> WT-F	GGTTCCATCACGGAGTTATATGG	55
<i>AtRAD51</i> WT-R	AGCCATGATATTCCCACCAATC	
<i>AtRAD51</i> MT-F	GAAACTTTTTGCTTTCCTTCTCTG	55
<i>AtRAD51</i> MT-R	ATATTGACCATCATACTCATTGC	
<b><i>AtDMC1</i> WT and T-DNA primers</b>		
<i>AtDMC1</i> WT-F	CAGAGGCATGAGATGAGATTTTACT	55
<i>AtDMC1</i> WT-R	GATTTGGTTAAAAGGATATGGCTTC	
<i>AtDMC1</i> MT-F	ATGCATACCAAGAAGGTATTAACCA	55
<i>AtDMC1</i> MT-R	AAGTTGTCTAAGCGTCAATTTGTTT	

**2.2.16.3 Ligation of the *CaMV35S::TaRAD51(D)::±GFP* and *CaMV35S::TaDMC1(D)::±GFP* constructs for transformation into *Arabidopsis rad51* and *dmc1* T-DNA mutants, respectively**

The *CaMV35S::TaRAD51(D)::±GFP* and *CaMV35S::TaDMC1(D)::±GFP* complementing constructs were generated as a translational fusion by ligating a DNA fragment comprising the promoter with or without the 5' end of GFP gene to *TaRAD51* and *TaDMC1* cDNA sequences. The *TaRAD51* and *TaDMC1* cDNA fragments (1.1kb) were generated from wild-type Highbury wheat spikelet cDNA by RT-PCR using primers which amplify all three homoeologues (Table 2.6) and contain restriction sites for insertion into the pGVT17 binary vector (Figure 2.2). Two different kinds of constructs were prepared for each gene using two sets of primers, one set (*Xba*+*Asc*) has no stop codon at the 3' end and so GFP in the binary vector pGVT17 is expressed by read-through (*CaMV35S::TaRAD51(D)::+GFP* and *CaMV35S::TaDMC1(D)::+GFP*) (Figure 2.9) and the second set



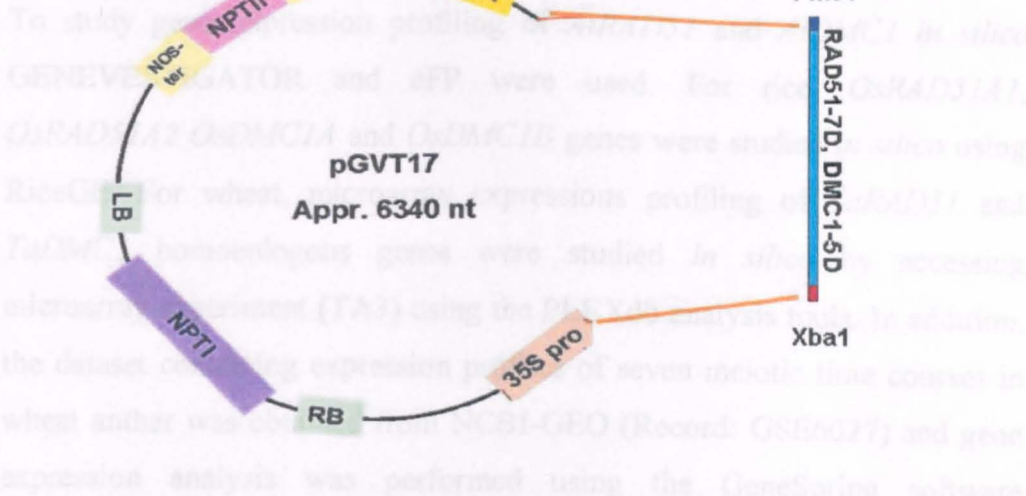
(*Xba*/*Pme*) has a stop codon at the 3' end and GFP is removed from the vector prior to insertion of the wheat gene. (*CaMV35S::TaRAD51(D)* and *CaMV35S::TaDMC1(D)*) (Figure 2.10). The amplified RT-PCR products with gene specific primers were sub-cloned into the pGEM T-Easy vector. The transformed *E. coli* colonies were then screened for the D-homoeologues of *TaRAD51* and *TaDMC1* cDNA clones by colony PCR using the D genome-specific primers (Table 2.6). These clones were confirmed by sequencing. After digestion with the *Xba*+*Asc* and *Xba*+*Pme* restriction enzymes, they were cloned into binary vector pGVT17 which has been restricted with the same set of enzymes as the cloned fragments. The constructs were then electroporated into the *Agrobacterium* EHA105 strain and the constructs were re-confirmed by colony PCR.



**Figure 2.9** Cloning of the *TaRAD51(D)* and *TaDMC1(D)* genes with fusion to *GFP5* into the pGVT17 binary vector.

2.2.17 Microarray methods

2.2.17.1 Microarray



**Figure 2.10** Cloning of the *TaRAD51(D)* and *TaDMC1(D)* genes into the pGVT17 binary vector without the GFP5 fusion.

2.2.17.2 Probe set identification

**Table 2.6** List of primers used for the isolation of ORF's and genome-specific clones of the *TaRAD51* and *TaDMC1* genes

Primer name	Sequence(5'-3')	Tm (°C)
<b><i>TaRAD51</i> gene primers (to amplify ORFs)</b>		
<i>TaRAD51</i> -Xba-F	ATCTAGAAAGGAGATATAACAATGTCGTCGGCGGCGGC	60
<i>TaRAD51</i> -Asc-R	AGGCGCGCCGATCCTTAACATCTGTGACGCCT	
<i>TaRAD51</i> -Pme-R	AGTTTAAACTCAATCCTTAACATCTGTGACGCCT	
<b><i>TaDMC1</i> gene primers (to amplify ORFs)</b>		
<i>TaDMC1</i> -Xba-F	ATCTAGAAAGGAGATATAACAATGGCGCCGTCCAAGCAG	60
<i>TaDMC1</i> -Asc-R	AGGCGCGCCGGTCTTTACATCCATCAATCCGCCTG	
<i>TaDMC1</i> -Pme-R	AGTTTAAACTCAGTCTTTACATCCATCAATCCGCCT	
<b>D genome-specific primers</b>		
<i>TaRAD51D</i> -F	TGGGGATACATCGTGTATTGGCCT	65
<i>TaRAD51D</i> -R	TGGTGGTCCAATATCACATAGGAG	
<i>TaDMC1D</i> -F	ATTAAGCGAAAGTCTGTTGTCC	65
<i>TaDMC1D</i> -R	CTCATAGGTGTAGGCACGAGCA	

## **2.2.17 Microarray methods**

### **2.2.17.1 Microarray data access**

To study gene expression profiling of *AtRAD51* and *AtDMC1* *in silico* GENEVESTIGATOR and eFP were used. For rice, *OsRAD51A1*, *OsRAD51A2* *OsDMC1A* and *OsDMC1B* genes were studied *in silico* using RiceGE. For wheat, microarray expressions profiling of *TaRAD51* and *TaDMC1* homoeologous genes were studied *in silico* by accessing microarray experiment (TA3) using the PLEXdb analysis tools. In addition, the dataset containing expression profiles of seven meiotic time courses in wheat anther was obtained from NCBI-GEO (Record: GSE6027) and gene expression analysis was performed using the GeneSpring software (Version:GX10, Agilent technologies, USA).

### **2.2.17.2 Probe set identification**

For Arabidopsis, the cDNA sequences of *AtRAD51* and *AtDMC1* were used in the query tool of PLEXdb (<http://www.plantgdb.org/cgi-bin/prj/PLEXdb/ProbeMatch.pl>) using Arabidopsis ATH1 22K. For wheat, after the cloning of *TaRAD51* and *TaDMC1* cDNA homoeologues were done, the cDNA homoeologue sequences were used in the query tool of PLEXdb using Wheat 61K gene chips to identify probe sets. For rice, the locus names (TIGR) of the *OsRAD51A1* (LOC\_11g40150), *OsRAD51A2* (LOC\_Os12g31370), *OsDMC1A* (LOC\_Os12g04980) and *OsDMC1B* (LOC\_Os11g04954) genes were used in the query tool in RiceGE (<http://signal.salk.edu/cgi-bin/RiceGE>) and the Os\_51K array was chosen.

### **2.2.17.3 BLAST for probe specificity**

The oligonucleotide sequences from each perfect match probe in each probe set for *RAD51* and *DMC1* examined were submitted to the NCBI database (<http://www.ncbi.nlm.nih.gov/BLAST/>). Using the BLAST tool, the number of perfect and partial sequence matches was recorded for each probe set for which sequence was available within the databases.

#### **2.2.17.4 GeneSpring Software**

Analyses of the wheat meiotic time course microarray datasets were primarily performed using GeneSpring. The data, comprising the expression profiles of seven meiotic time courses in the wheat anther, was obtained from NCBI-GEO (Record: GSE6027) loaded and normalized in GeneSpring using the default normalization methods. These included setting any signal values below 0.01 to 0.01, total chip normalization to the 50th percentile, and normalization of each gene to the median based on its measured levels of expression. A major aim of this work was to determine the expression profiling of the *TaRAD51* and *TaDMC1* genes expressed during seven meiotic time courses in wheat anther. The gene expression profiles of both *TaRAD51* and *TaDMC1* genes were then studied after entering the probe-set id numbers of the corresponding genes.

## CHAPTER 3: MICROARRAY EXPRESSION

### PROFILING OF *RAD51* AND *DMC1* IN ARABIDOPSIS, RICE AND WHEAT

#### 3.1 Introduction

The patterns of gene expression and simultaneous expression assessment of thousands of genes at a given time can be determined at the genome level using microarray technologies. In plants, microarrays have become a widely-used tool for transcriptome analysis. Due to the advancements of genome technologies many kinds of microarray technology platforms that can measure global gene expression in plants have arrived (Galbraith, 2006). Essentially, microarrays consist of an immobilized array of oligonucleotides (probes) which are hybridized *in situ* using fluorescently-labelled sequences which are prepared by reverse transcription of polyadenylated transcripts. The process of measurement of gene expression by microarray starts by extraction of mRNA from the cells or tissues of interest and hybridizing the same to the 25-mer probes on the microarray, for Affymetrix arrays (Affymetrix, 2001). This is followed by imaging of the hybridized array and computational analysis using microarray analysis software. Through microarrays, the RNA transcript level can be measured either in relative terms using glass slide microarrays (Schena *et al.*, 1995) or in absolute terms using Affymetrix GeneChip arrays (Lipshutz *et al.*, 1999).

Each gene is represented on an Affymetrix GeneChip array by a probe set. The probe set on the microarray consists of a perfect match probe (PM), with a 25-base sequence identical to the gene of interest, and a mismatch probe (MM), whose sequence is the same as the perfect match except for a mismatch at thirteen position, where the base is set to the complementary of the perfect match. The mismatch probe was introduced by Affymetrix as a measure of

cross-hybridisation during microarray experiments to assess potential background levels of 'noise'. This probe pairing design helps to identify and deduct the non-specific hybridization and background signal for microarray analysis. Probe sets ideally represent unique regions of the gene allowing detection for specific and individual gene transcripts, even among genes from the same gene family. In practice, this can be difficult to achieve without complete genome sequence information. Each array also contains probe sets for controls that are spiked into the sample prior to hybridization. Determination of gene expression using microarray consists of measurement of temporal and spatial expression patterns that can be associated with a specific phenotype or in response to a specific stimulus or in a specific tissue in a microarray experiment. In practice, the design of the experiment and use of appropriate materials is more important than the technical aspects of carrying out the experiment itself.

The experimental data from microarray experiments can be deposited in high-throughput gene expression / hybridization data repositories, such as the NCBI-Gene Expression Omnibus (GEO). However, this repository only make the data sets available to public users, but they do not readily allow researchers to search for specific genes of interest based on sequence information or community annotation, nor do they allow flexible on-line analysis that provides links between probe set expression and probe set annotation. Online microarray database analysis tools will help to resolve this problem and can be used to identify the expression profile and function of an unknown gene by accessing a large collection of expression data from mutants, tissues or treatments using a variety of gene expression databases. Some of the examples of the databases which provide species-specific resources for *Arabidopsis*, are TAIR (Rhee *et al.*, 2003), GENEVESTIGATOR (Zimmerman *et al.*, 2004) and NASC Arrays (Craigon *et al.*, 2004). TAIR provides detailed annotation information for *Arabidopsis* genes as well as tools for microarray data analysis. GENEVESTIGATOR has an extensive set of on-line tools for analysis and

visualization of Arabidopsis datasets. NASC Arrays uses all of the GeneChip data from their facility and other projects publicly available for download for use in relational databases and for further expression pattern analysis. The Rice Expression Database (RED) and the associated Rice Pipeline (Yazaki *et al.*, 2004) project compile genomics data (genome sequences, full-length cDNAs, gene expression profiles, mutant lines, *cis* elements) into one data source with full annotation. For barley there are many reference experiments using the barley 22K GeneChip (Close *et al.*, 2004) available through the BarleyBase expression database (Lishuang *et al.*, 2005) and analysis tools for gene expression are also available through Genevestigator. BarleyBase has expanded to PLEXdb (Plant expression database), with the addition of expression data from several plant species such as Arabidopsis, barley, grape, maize, *Medicago*, soybean, *Fusarium*, rice, tomato and wheat. Expression data experiments for soybean generated using a spotted cDNA arrays are available through the Soybean Genomics and Microarray Database (SGMD) (Alkharouf and Matthews, 2004). Likewise large quantities of expression data for the Solanaceae, including potato and tomato, are available through the Solanaceae Gene Expression Database, and the tomato gene expression data are also available through a tomato expression database (Table 3.1).

**Table 3.1** Publicly available plant microarray expression databases and their URLs.

<b>Plant species</b>	<b>Database</b>	<b>URL address</b>
Arabidopsis	NASC Arrays	<a href="http://affymetrix.Arabidopsis.info/narrays/experimentbrowse.pl">http://affymetrix.Arabidopsis.info/narrays/experimentbrowse.pl</a>
Arabidopsis	TAIR	<a href="http://arabidopsis.org">http://arabidopsis.org</a>
Arabidopsis	GENEVESTIGATOR	<a href="http://www.genevestigator.ethz.ch/">http://www.genevestigator.ethz.ch/</a>
BAR	Poplar Expression database	<a href="http://BAR.utoronto.ca">http://BAR.utoronto.ca</a>
Barley	Barleybase	<a href="http://www.barleybase.org">http://www.barleybase.org</a>
Generic	Gene Expression Omnibus (GEO)	<a href="http://www.ncbi.nlm.nih.gov/projects/geo">http://www.ncbi.nlm.nih.gov/projects/geo</a>
Generic	PLEXdb	<a href="http://www.barleybase.org/plexdb/html/index.php">http://www.barleybase.org/plexdb/html/index.php</a>
Generic	ArrayExpress	<a href="http://www.ebi.ac.uk/arrayexpress">http://www.ebi.ac.uk/arrayexpress</a>
Rice	NSF Rice Oligonucleotide Array Project and Database	<a href="http://www.ricearray.org">http://www.ricearray.org</a>
Rice	Virtual Center for Cellular Expression Profiling in Rice	<a href="http://130.132.8.83/rc">http://130.132.8.83/rc</a>
Rice	Rice Expression Database	<a href="http://red.dna.affrc.go.jp/RED">http://red.dna.affrc.go.jp/RED</a>
Solanaceae	Solanaceae Gene Expression Database	<a href="http://www.tigr.org/tdb/potato">http://www.tigr.org/tdb/potato</a>
Soybean	Soybean Genomics and Microarray Database	<a href="http://psi081.ba.ars.usda.gov/SMD/default.htm">http://psi081.ba.ars.usda.gov/SMD/default.htm</a>
Tomato	Tomato Expression Database	<a href="http://ted.bti.cornell.edu">http://ted.bti.cornell.edu</a>



Comprehensive Affymetrix GeneChip platforms have now been developed for wheat based on extensive EST collections. PLEXdb supports all currently available array platforms and it provides an integrated web interface for functional interpretation of highly parallel microarray experiments. It can also provide gene expression data in the form of expression graphs, hierarchical clustering and heat maps for a variety of crop and model plant species, from a large array of experimental or field conditions. The array data in turn are interlinked to analytical and biological functions (e.g., Gene and Plant Ontologies, BLAST, spliced alignment, multiple alignment, regulatory motif identification, and expression analysis), allowing anyone to access and analyze comparative expression experiments in conjunction with their own data.

PLEXdb is complementary but not repetitive in comparison to other microarray databases such as NCBI - GEO (Edgar *et al.*, 2002; Barrett *et al.*, 2005), Stanford microarray Database (Gollub *et al.*, 2003; Ball *et al.*, 2005) and ArrayExpress (Brazma *et al.*, 2003). PLEXdb supports all of the available plant and pathogen GeneChip arrays including, Affymetrix 57K rice, 61K wheat, 22K barley1, 18K first-generation maize, 8K sugar cane, *Fusarium graminearum*, 61K soybean/*Phytophthora*/cyst nematode, 16K grape as well as cotton, poplar, citrus, tomato, *Medicago*, and the full-genome maize, with complete annotation and analysis support.

GENEVISTIGATOR is both a database and web-browser data mining interface for facilitating expression analysis of Affymetrix Gene Chip data for Arabidopsis and other species, such as barley, mouse, yeast, *E. coli* and human. It comprises of a number of online querying and analysis tools to identify gene function. GENEVESTIGATOR specifically allows the data to be presented in the terms of plant development, plant organ, and environmental conditions, both for individual genes or for families of genes. The main objective of the software is to assign contextual information to gene expression data and allows the design of new experiments and gene functional discovery. The

GENEVESTIGATOR allows two types of queries in its software structure: The first one is a gene-centric approach reporting signal intensity values for individual genes and the second one is a genome-centric approach providing lists of genes fulfilling chosen criteria. The results obtained from any tool are based on all available signal intensity values and the corresponding annotations. The different functional tools incorporated in the GENEVESTIGATOR include Digital Northern, an online analysis tool for retrieving the signal intensity values of a gene of interest for a chosen selection of GeneChip experiments. Up to 10 probe sets can be queried simultaneously and will be displayed in several colors, shapes, with filling revealing both signal intensity values and 'present' call (closed symbols) and 'absent' call (open symbols) information. The Gene Correlator is an analysis tool for comparing the signal intensities of two genes throughout all chosen experiments. Gene Atlas is a tool for providing the average signal intensity values of a gene of interest in all organs or tissues annotated in the database.

Expression analysis tools at the BAR (The Bio-Array Resource for Arabidopsis Functional Genomics) include the Expression Angler, the electronic Fluorescent Pictograph (eFP) Browser and e-Northerns with Expression Browser tools. The Arabidopsis eFP (<http://www.bar.utoronto.ca/>) browser is an online microarray analysis tool for interpreting and exploring large-scale biological datasets (Winter *et al.*, 2007). The eFP Browser engine allows the input of large-scale datasets onto pictographic representations of the experimental samples used to generate the data sets. The gene of interest is queried in the form of the AGI ID (Arabidopsis Genome Initiative identifier) and upon submission, the plant tissues are coloured according to the expression level of the gene of interest in a particular tissue under a particular treatment condition. The eFP tool is a quick and easy means of identifying tissues where there is differential expression and is particularly useful when exploring gene families to facilitate hypothesis generation. There are three different modes in eFP for the functional analysis of the gene of interest. In "Absolute," mode the

expression level for a gene of interest in each tissue is compared directly to the highest signal recorded for the given gene, with low levels of expression are indicated as yellow and high levels are coloured red. The “Relative” mode displays the ratio of a tissue's expression level to appropriate control signal – typically the median or mock treatment – for its group. The output in relative mode has tissues with expression levels above the control signal value coloured between yellow and red, and expression levels below the control signal value coloured between yellow and blue. In the “Compare” mode expression of two genes can be compared directly. The eFP browser accepts two gene identifiers (two genes of interest) as input and then compares the primary relative expression level of one gene to another in each of the tissue using the same colouring scheme as for the Absolute mode. The compare mode is useful for identifying tissues in which one gene is more abundantly expressed relative to another.

Rice GE is a Rice Functional Genomic Express Database tool developed by the Salk Institute (<http://signal.salk.edu/cgi-bin/RiceGE>).

In this current chapter, transcriptional expression profiling pattern of the *AtRAD51*, *AtDMC1* genes were studied *in silico* using GENEVESTIGATOR and Arabidopsis eFP browser. The *OsDMC1* and *OsRAD51* genes were studied using Rice GE. The studies of both spatial and temporal transcriptional expression patterns of the *TaDMC1* and *TaRAD51* genes were studied *in silico* using online microarray database tool such as PLEXdb and the GeneSpring software respectively. However it is well known fact that meaningful comparative expression analyses using microarray platforms based solely on EST collections can be difficult because of the frequent and confounding presence of multiple splice forms, paralogues and orthologues, as well as, in the case of polyploids, homoeologues with near-identical sequence (Poole *et al.*, 2007). So care must be taken when performing comparative transcript analysis of these genes. Because the 61,115 probe sets on the Wheat GeneChip reflect

the complete collection of publicly available wheat ESTs at the time of design of the chip, this dataset serves as an 'expression atlas' for hexaploid wheat, rather than comprehensive transcriptome coverage.

### 3.2 Results

#### 3.2.1 Identification of the probe sets corresponding to the *RAD51* and *DMC1* genes of Arabidopsis, rice and wheat

The probe sets and expressed probe sequences (oligonucleotides) corresponding to the probe sets of *AtRAD51*, *AtDMC1*, *OsRAD51A1*, *OsRAD51A2*, *OsDMC1A*, *OsDMC1B*, *TaRAD51* and *TaDMC1* were obtained from PLEXdb (Table 3.2).

**Table 3.2** Probe set identification numbers of *RAD51* and *DMC1* genes of Arabidopsis, rice and wheat.

Gene name	AGI code/ Accession no	Probe set id
<i>AtRAD51</i>	At5g20850	246132_at
<i>AtDMC1</i>	At3g22880	256832_at
<i>OsRAD51A1</i>	LOC_Os11g40150	Os.57536.1.S1_at
<i>OsRAD51A2</i>	LOC_Os12g31370	Os.19630.1.S1_at
<i>OsDMC1A</i>	LOC_Os12g04980	Os.13051.1.S1_at
<i>OsDMC1B</i>	LOC_Os11g04954	Os.22122.1.S1_at
<i>TaRAD51-7A</i>	FJ594479*	Ta.7197.1.A1_at
<i>TaRAD51-7B</i>	FJ594480*	Ta.7197.1.A1_at
<i>TaRAD51-7D</i>	FJ594481*	Ta.7197.1.A1_at
<i>TaDMC1-5A</i>	FJ594476*	Ta.30833.1.S1_at
<i>TaDMC1-5B</i>	FJ594477*	Ta.30833.1.S1_at
<i>TaDMC1-5D</i>	FJ594478*	Ta.30833.1.S1_at

\*submitted by Devisetty *et al*, (24/12/2008)

### **3.2.2 Gene specificity of probe sequences of *RAD51* and *DMC1* in *Arabidopsis*, rice and wheat**

Before subjecting the probe sets for expression analysis, the individual oligonucleotide probe sequences corresponding to each gene were checked by BLASTN to confirm their target gene specificity. The BLASTN examination of probe sequences representing *AtRAD51*, *AtDMC1*, *OsRAD51A1*, *OsRAD51A2*, *OsDMC1A* and *OsDMC1B* by BLASTN revealed that the oligonucleotides are represented only by the target gene, with little or no chance of cross-hybridization to other genes (Table 3.3, 3.4, 3.5, 3.6, 3.7 and 3.8). All the three cDNA homoeologues of *TaRAD51* as well as all three cDNA homoeologues of the *TaDMC1* were being represented by only a single probe set on the wheat Affymetrix chip. The examination of individual oligonucleotides within the probe sets for both *TaRAD51* and *TaDMC1* indicated that they match one or two or all cDNA homoeologues (Table 3.9 and 3.10).

**Table 3.3** BLASTN search results of the oligonucleotides (probes) in *AtRAD51* probe set (246132\_at)

Gene and Probe ID	Probe sequence (5'-3')	BlastN hit	E-value	Length
<i>AtRAD51</i> At5g20850 246132_at	GGAACATTCAGGCCACAAAGACTCT	At5g22880	1.20E-07	25
		At5g19400	0.76	15
	AAACGTTGCCTATGCGAGGGCGTAT	At5g20850	1.20E-07	25
	ACAGATCATCAGTCAAGGCTTTTGC	At5g20850	1.20E-07	25
		At1g66360	0.75	15
		At5g53060	0.76	15
	AACAAGGTTTGCTCTCCTGATTGTC	At5g20850	1.20E-07	25
		At3g02950	0.75	15
		At2g06000	0.76	15
	TGATTGTCGATAGTGCTACCGCTCT	At5g20850	1.20E-07	25
		At1g42180	0.74	15
	CTACCGCTCTCTACAGAACAGATTT	At5g20850	1.20E-07	25
		At2g28370	0.75	15
	GAGCTTTCGGCTCGACAAATGCATC	At5g20850	1.20E-07	25
	TGGTTCAGCTCTTTTGCTGGTCCC	At5g20850	1.20E-07	25
		At2g01505	0.76	15
		At3g55320	0.76	15
	TGCTGGTCCCAATTTAAGCCGATT	At5g20850	1.20E-07	25
	ATATCATGGCTCATGCCACCACAAC	At5g20850	1.20E-07	25
		At4g01910	0.76	15
		At2g16390	0.76	15
	TGATAAGCTCGCCATGTTTGCCAGA	At5g20850	1.20E-07	25
		At1g24764	0.76	15

**Table 3.4** BLASTN search results of the oligonucleotides (probes) in *AtDMC1* probe set (256832\_at)

Gene and Probe ID	Probe sequence (5'-3')	BlastN hit	E-value	Length
<i>AtDMC1</i> At3g22880 256832_at	CACTACGCAGCTACCTACAAACATG	At3g22880	1.20E-07	25
	GGAACCTTCCGTCCTGATAGGATTG	At3g22880	1.20E-07	25
	ATTTATGCTCGTGCTTATACCTATG	At3g22880	1.20E-07	25
	TCAGTACAACCTTGCTTCTTGGCCTT	At3g22880	1.20E-07	25
	TTGACTCGATCATTGCTTTATTCCG	At3g43128	0.76	15
		At3g22880	1.20E-07	25
		At4g23630	0.16	20
		At3g27500	0.16	16
	ACTAGCTCAGATGCTTTCCAGGCTA	At3g22880	1.20E-07	25
		At2g39240	0.035	17
		At4g16330	0.75	15
	GAGGAGTTCAACGTTGCTGTCTACA	At3g22880	1.20E-07	25
		At4g19985	0.16	16
	GTGGTCATGTACTAGCTCACGCAGC	At3g22880	1.20E-07	25
		At5g20770	0.74	15
	GTCTGCAAAGTCTACGATGCTCCGA	At3g22880	1.20E-07	25
	TGCTCCGAATCTCGCTGAAGCTGAA	At3g22880	1.20E-07	25
	AGCTGAAGCCTTTCCAGATTACTCA	At3g22880	0.0013	25



**Table 3.5** BLASTN search results of the oligonucleotides (probes) in *OsRAD51A1* probe set (Os.57536.1.S1\_at)

Gene and Probe ID	Probe sequence (5'-3')	BlastN hit	E-value	Length
<i>OsRAD51A1</i> LOC_Os11g40150 Os.57536.1.S1_at	AAGCTCTCCATGCTTGGCTGAAGCC	LOC_Os11g40150	2.30E-07	25
		LOC_Os10g42660	0.31	16
		LOC_Os04g15990	1.4	15
	GCCGAAGCAAGGTTTCAGATAGCAT	LOC_Os11g40150	2.30E-07	25
		LOC_Os04g17900	1.5	15
		LOC_Os07g31460	1.5	15
	TCTCTGCTCAGTGCTAAGCGCTAAC	LOC_Os11g40150	2.30E-07	25
		LOC_Os03g02640	1.4	15
		LOC_Os10g43075	1.5	15
	TAAGCGCTAACTGGCAACAGCTTCC	LOC_Os11g40150	2.30E-07	25
	AGTTCCCACATGTGTTGCTCTCCTG	LOC_Os11g40150	2.30E-07	25
	TGATTGCGCTACCACCCCTGAAGGG	LOC_Os11g40150	2.30E-07	25
		LOC_Os03g08500	1.4	15
		LOC_Os04g16100	1.5	15
	AACTGATGTGATGCTCCTGTTCTGA	LOC_Os11g40150	2.30E-07	25
		LOC_Os06g17930	1.5	15
		LOC_Os07g10150	1.5	15
	GTTCTGAATCCGACAACCACTTATT	LOC_Os11g40150	2.30E-07	25
		LOC_Os06g39290	1.4	15
	GTGCACATCTGTGTCAGTGTGTTAG	LOC_Os11g40150	2.30E-07	25
	GTATAAAACGTTGTCTGCCCTGCAA	LOC_Os11g40150	2.30E-07	25
		LOC_Os06g45910	1.5	15
	TGCCCTGCAAGGCTGCAATGCATGT	LOC_Os11g40150	2.30E-07	25
		LOC_Os03g06320	0.064	17
		LOC_Os03g52180	0.066	17

**Table 3.6** BLASTN search results of the oligonucleotides (probes) in *OsRAD51A2* probe set (Os.19630.1.S1\_at)

Gene and Probe ID	Probe sequence (5'-3')	BlastN hit	E-value	Length
<i>OsRAD51A2</i> LOC_Os12g31370 Os.19630.1.S1_at	CTGGCCAAGTTTCTTAGGAGCCTTC	LOC_Os12g31370	2.30E-07	25
		LOC_Os02g03060	1.4	15
		LOC_Os02g03060	1.4	15
	GTGCTGCAATGTTTGGGCCACAGAT	LOC_Os12g31370	2.30E-07	25
		LOC_Os02g25260	0.067	17
	GCCACAGATCAAACCTATCGGCGGG	LOC_Os12g31370	2.30E-07	25
		LOC_Os04g57220	0.31	16
		LOC_Os02g58550	1.4	15
	TATCGGCGGGAACATCATGGCTCAT	LOC_Os12g31370	2.30E-07	25
		LOC_Os11g40150	0.31	24
		LOC_Os09g39680	1.5	15
	GTAAGGTAGTAAGCTCTCCCTGTCT	LOC_Os12g31370	2.30E-07	25
	GATATCACCGGAAGGCGTCACAGAT	LOC_Os12g31370	2.30E-07	25
		LOC_Os05g34140	1.5	15
	TTGCAATCCCAGTGTGGTGGACGCT	LOC_Os12g31370	2.30E-07	25
		LOC_Os04g24490	1.5	15
	ATTGATGTAGGCCATCCTGTTCAAC	LOC_Os12g31370	2.30E-07	25
	TAGCGCCCTGGAATCCTGGATGGAA	LOC_Os12g31370	2.30E-07	25
	GGGTGTACAGCAAGTCGCCACAATT	LOC_Os12g31370	2.30E-07	25
		LOC_Os01g54960	1.4	15
		LOC_Os06g03710	1.5	15
	CTTGGCACCAGCGATGTTGTTGTTTC	LOC_Os12g31370	2.30E-07	25
		LOC_Os02g51060	0.32	16

**Table 3.7** BLASTN search results of the oligonucleotides (probes) in *OsDMC1A* probe set (Os.13051.1.S1\_at).

Gene and Probe ID	Probe sequence (5'-3')	BlastN hit	E-value	Length
<i>OsDMC1A</i> LOC_Os12g04980 Os.13051.1.S1_at	TAACTGACCTGAAAAAACCAGCGGG	LOC_Os12g04980	2.30E-07	25
	CCATATAGGTTTGGTACATTGACTA	LOC_Os12g04980	2.30E-07	25
		LOC_Os01g52760	1.4	1
	CATTGACTAACTACTACTGCTACT	LOC_Os12g04980	2.30E-07	25
		LOC_Os02g35180	0.31	20
		LOC_Os06g08130	1.4	15
	TACTACTGCTACTTTGCCACTGTCG	LOC_Os12g04980	2.30E-07	25
		LOC_Os02g13870	1.4	15
		LOC_Os01g13800	1.5	15
	GCTACTTTGCCACTGTCGAGGAAAT	LOC_Os12g04980	2.30E-07	25
		LOC_Os03g15760	1.4	15
	GCCACTGTCGAGGAAATATGCAACC	LOC_Os12g04980	2.30E-07	25
		LOC_Os03g58910	1.5	15
		LOC_Os03g58910	1.5	15
	ATATGCAACCTCATTTATCCAGACG	LOC_Os12g04980	2.30E-07	25
		LOC_Os11g04954	0.066	21
		LOC_Os08g39210	1.5	15
	TGCAACCTCATTTATCCAGACGATT	LOC_Os12g04980	2.30E-07	25
		LOC_Os11g04954	0.066	21
		LOC_Os08g39210	1.5	15
	ATCCAGACGATTATACCTTAAAATG	LOC_Os12g04980	2.30E-07	25
	ATGGGTATTTTCTATGCTTATGAG	LOC_Os12g04980	2.30E-07	25
	ATGCTTATGAGATCAACAGTTGTAC	LOC_Os12g04980	2.30E-07	25
		LOC_Os11g04954	0.066	21
		LOC_Os01g12650	1.4	15

- indicates potentially cross-hybridizing probes

**Table 3.8** BLASTN search results of the oligonucleotides (probes) in *OsDMC1B* probe set (Os.22122.1.S1\_at).

Gene and Probe ID	Probe sequence (5'-3')	BlastN hit	E-value	Length
<i>OsDMC1B</i> LOC_Os11g04954 Os.22122.1.S1_at	G TTCATAACTGACCCGAAGAAACCT	LOC_Os11g04954	2.30E-07	25
		LOC_Os04g27630	0.32	16
		LOC_Os08g16190	1.5	15
	C CGAAGAAACCTGCGGGAGGCCACG	LOC_Os11g04954	2.30E-07	25
		LOC_Os02g40520	1.4	15
		LOC_Os05g27970	1.4	15
	C CATATAGGTTAGGTACATTGATAA T GCTATGTTTTTGCCACTGCTGAGG	LOC_Os11g04954	2.30E-07	25
		LOC_Os11g04954	2.30E-07	25
		LOC_Os05g26290	0.067	17
	G TTTTTGCCACTGCTGAGGAAAGAG	LOC_Os05g26170	0.067	17
		LOC_Os11g04954	2.30E-07	25
		LOC_Os05g41510	0.32	16
	A GAGGCAACCTCATTATCAAGACG	LOC_Os06g02590	1.5	15
		LOC_Os11g04954	2.30E-07	25
		LOC_Os12g04980	0.066	21
	G ACGGTTATATCTCACATTGGGTAT T GAGTTATGTGATCAACAGTTGTAC	LOC_Os11g35860	1.4	15
		LOC_Os11g04954	2.30E-07	25
		LOC_Os11g04954	2.30E-07	25
	G TTGTACTGAAACAAGTGCTAGTAT	LOC_Os12g04980	0.066	21
		LOC_Os01g33370	1.4	15
		LOC_Os11g04954	2.30E-07	25
	A GTGCTAGTATTGTCTTAAGTGAAA	LOC_Os12g04980	1.4	15
		LOC_Os01g54784	1.5	15
		LOC_Os11g04954	2.30E-07	25
	T ATAAGAGGTTAACATGTTGGTTTC	LOC_Os12g04980	0.0026	19
		LOC_Os11g04954	2.30E-07	25

- indicates cross-hybridizing probes

**Table 3.9** BLASTN search results of the oligonucleotides (probes) in *TaRAD51* probe set (Ta.7197.1.A1\_at).

No	Probe set i.d	Probe sequence	No. of nucleotides in oligos matching homoeologues			Overall genome assignment
			A	B	D	
1	Ta.7197.1.A1_at	GATCTGTAAGGTGGTGAGCTCTCCC	25	25	25	A,B,D
2	Ta.7197.1.A1_at	TCCCTGCCTGGCTGAAGCTGAAGCA	25	25	25	A,B,D
3	Ta.7197.1.A1_at	GATATCACCTGAAGGCGTCACAGAT	25	25	25	A,B,D
4	Ta.7197.1.A1_at	AGCCTGCTGTTTGCATCTTCAGCAA	25	25	25	A,B,D
5	Ta.7197.1.A1_at	AACAGATGCTAGCTTGTGCCGATGA	25	25	25	A,B,D
6	Ta.7197.1.A1_at	AATGTTGTACAGCTAATCTCCCAT	25	25	25	A,B,D
7	Ta.7197.1.A1_at	AATCTCCCATAGTTTGCTTGTCTA	25	25	25	A,B,D
8	Ta.7197.1.A1_at	GCTTGTCTATCTTTTTCAGTGGTT	25	23	23	A
9	Ta.7197.1.A1_at	ATGTAGTTGCTCACTTGTGTTCAAA	25	25	25	A,B,D
10	Ta.7197.1.A1_at	AAATCCTTGGTCTTCTAGGACATTG	25	23	23	A
11	Ta.7197.1.A1_at	GGACATTGACATTGTCTCCTATGTG	25	24	24	A

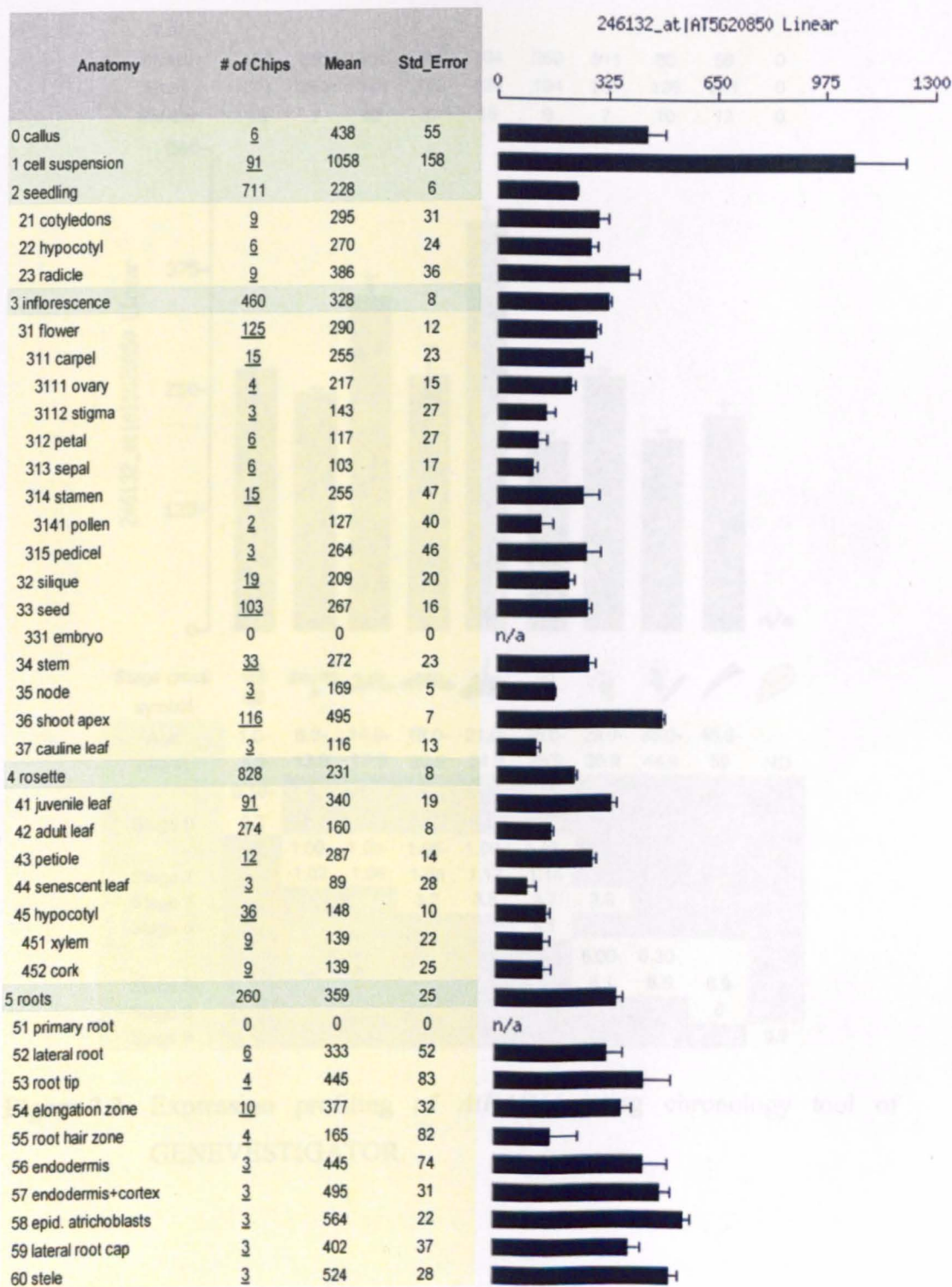
**Table 3.10** BLASTN search results of the oligonucleotides (probes) in *TaDMC1* probe set (Ta.7197.1.A1\_at).

No	Probe set i.d	Probe sequence (5'-3')	No. of nucleotides in oligos matching homoeologues			Overall genome assignment
			A	B	D	
1	Ta.30833.1.S1_at	TGGTATGTTCACTGACCCCAAA	25	25	25	A, B, D
2	Ta.30833.1.S1_at	TCTGCAAGATCTTTGACGCCCTAA	25	25	25	A, B, D
3	Ta.30833.1.S1_at	GGAGAAGCTGTTTTCCAGATCACAA	25	25	24	A, B
4	Ta.30833.1.S1_at	GACTGATTGTTCATCCAGGGCGCTT	25	25	25	A, B, D
5	Ta.30833.1.S1_at	GGCGCTTCCTGTCTTAAACTGTTGA	25	24	25	A, D
6	Ta.30833.1.S1_at	AAAGGCTACTGCACCGTTACTCTGT	25	25	25	A, B, D
7	Ta.30833.1.S1_at	GTTACTCTGTTTTGCCACTGCAGA	25	25	24	D
8	Ta.30833.1.S1_at	AATTGGGCAACCTCTATGCTGTTTA	25	25	25	A, B, D
9	Ta.30833.1.S1_at	TATGCTGTTTATACCTCCTGTTGGG	25	25	25	A, B, D
10	Ta.30833.1.S1_at	TGGGTGTTCCGTAAGCTTGTAATGC	25	25	24	D

### **3.2.3 Transcriptional expression profiling of the *RAD51* gene in Arabidopsis, rice and wheat using microarray databases**

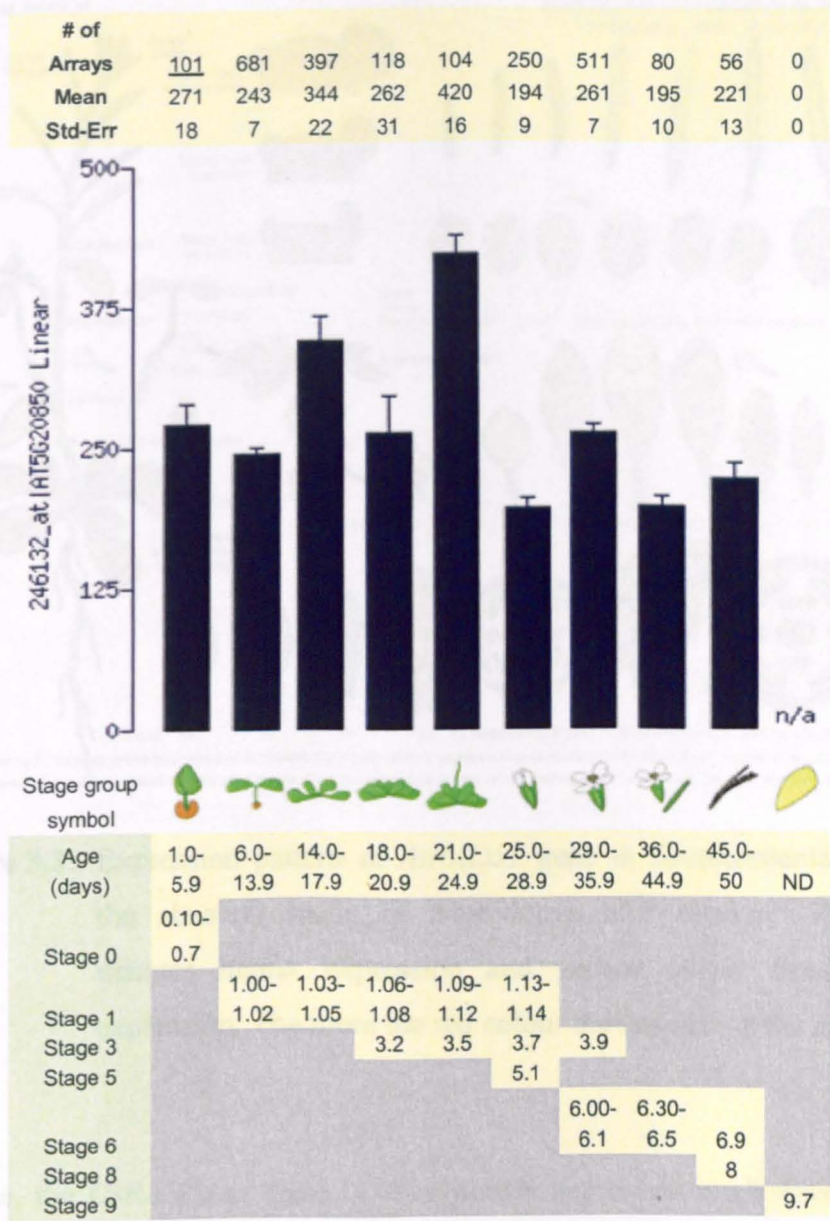
To investigate the expression of *RAD51* gene in Arabidopsis, rice and wheat publicly available microarray data were analyzed. The Gene Atlas tool of GENEVESTIGATOR showed that the maximum level of expression for *AtRAD51* is in cell suspension at signal intensity of 1038. There is a significant level of expression of *AtRAD51* in other vegetative tissues such as shoot apex and root tissue (Figure 3.1). Analysis of *AtRAD51* gene expression using the gene chronologer tool showed that the strongest gene expression was found to be at sub-stages: 14.0 to 17.9 days old seedling (rosette growth) and 21.0 to 24.9 days old seedlings (inflorescence emergence) (Figure 3.2). Arabidopsis eFP browser output for *AtRAD51* gene revealed the highest level of expression in the datasets to be in the shoot apex inflorescence, shoot apex (transition), shoot apex vegetative and vegetative rosette; weaker expression in flower stage 9, seed stage 10 and low expression levels in leaf stage 2 and 6, leaf cauline, leaf senescence, and dry seed (Figure 3.3).



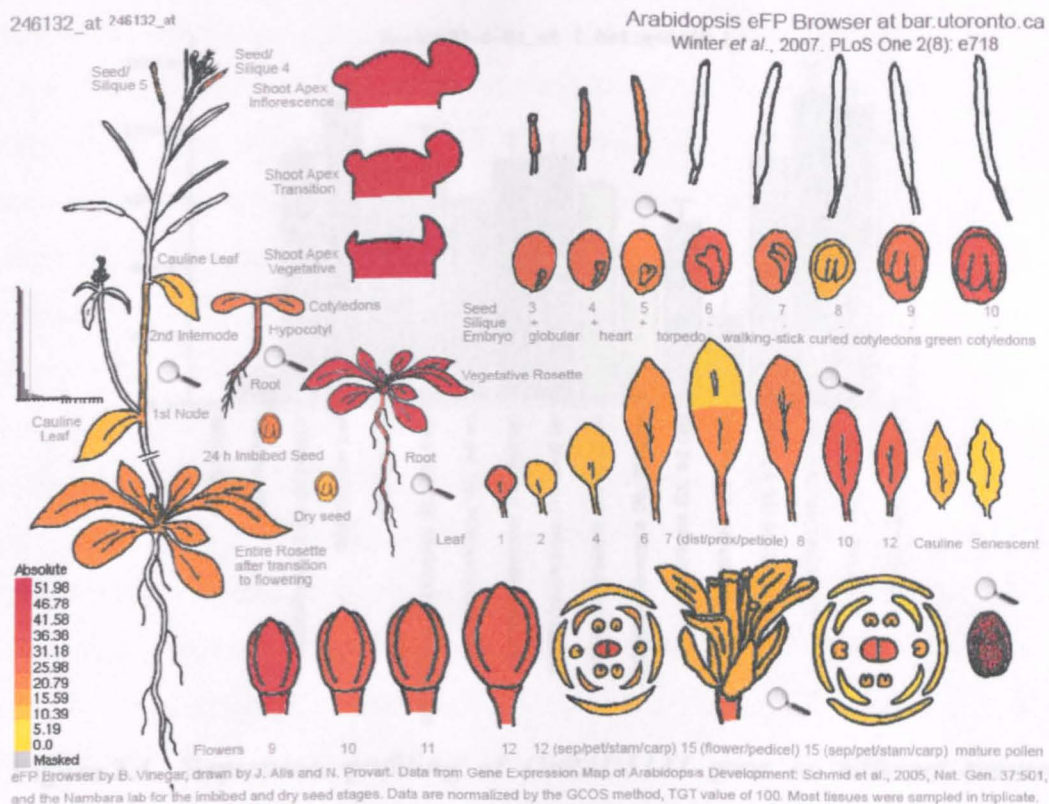


**Figure 3.1** Expression profiling of *AtRAD51* using gene atlas tool of GENEVESTIGATOR



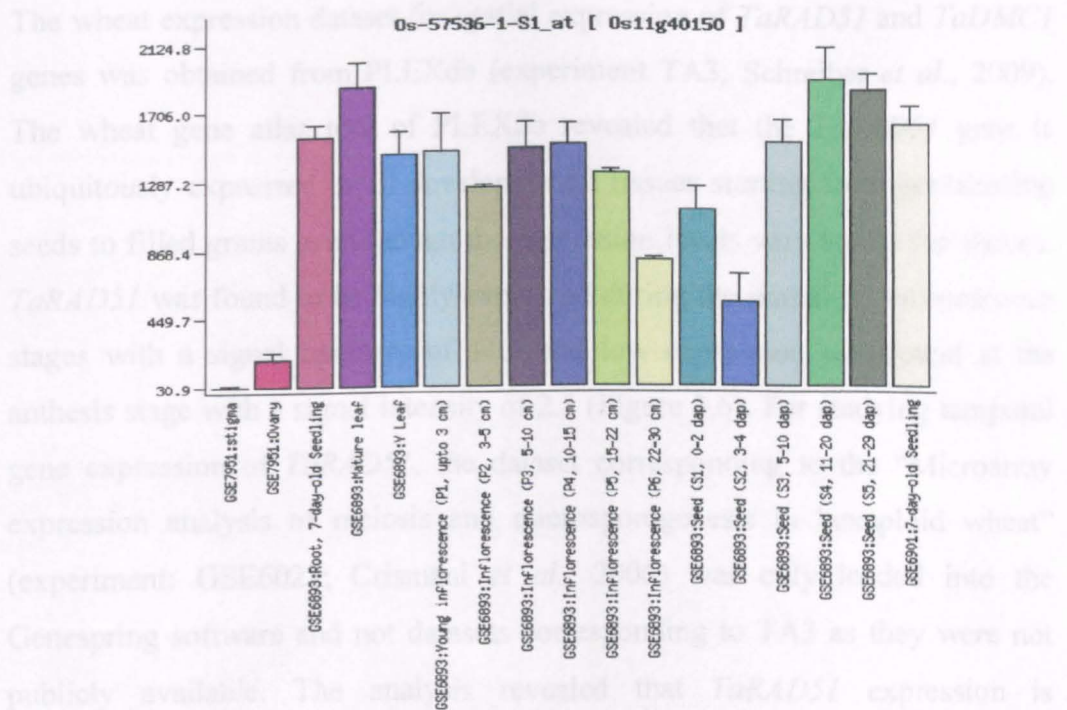


**Figure 3.2** Expression profiling of *AtRAD51* using chronology tool of GENEVESTIGATOR.

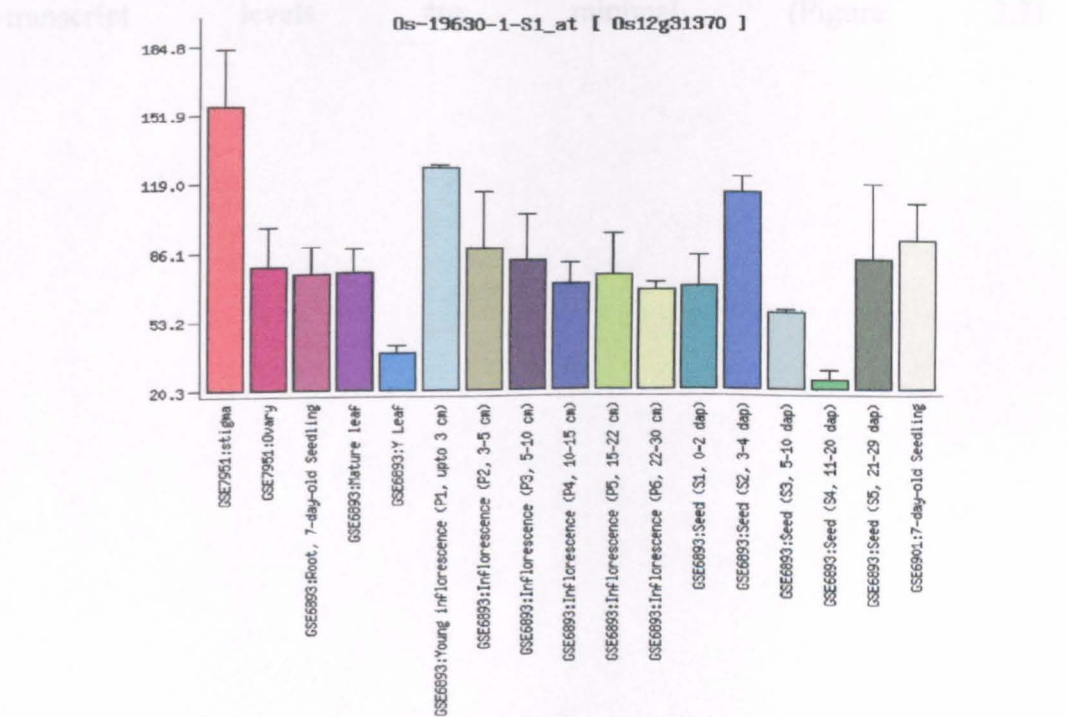


**Figure 3.3** Expression pattern of *AtRAD51* gene in developmental menu of the absolute mode of Arabidopsis eFP browser. Red colour denotes higher expression and yellow colour denotes lower expression. The more the red colour the stronger is the expression.

In rice, the *OsRAD51A1* gene is ubiquitously expressed in all developmental tissues analyzed. Higher expression levels were detected in many tissues with the mean signal intensity values ranging from 1,285 to 1900 and lower levels were detected in the stigma and ovary (Figure 3.4). On the other hand, *OsRAD51A2* showed lower expression levels in every tissue analyzed with the mean signal intensity values not more than 152. However, the highest level of signal intensity was detected in the stigma with the mean signal intensity at 152 (Figure 3.5; note that the Y-axis scale for Figure 3.4 and 3.5 are different).



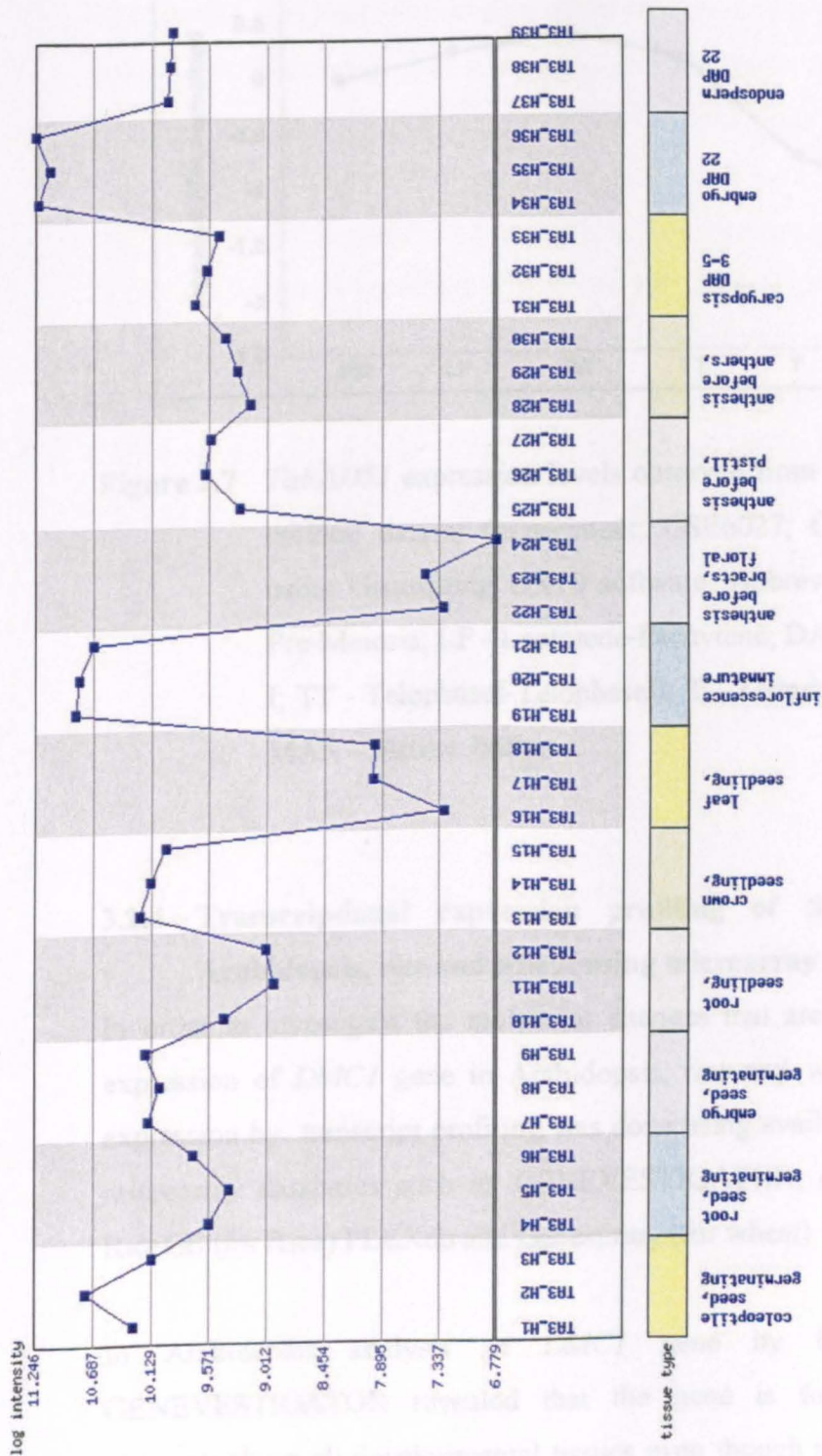
**Figure 3.4** Expression profiling of *OsRAD51A1* gene in different tissues using gene atlas tool in Rice GE database.



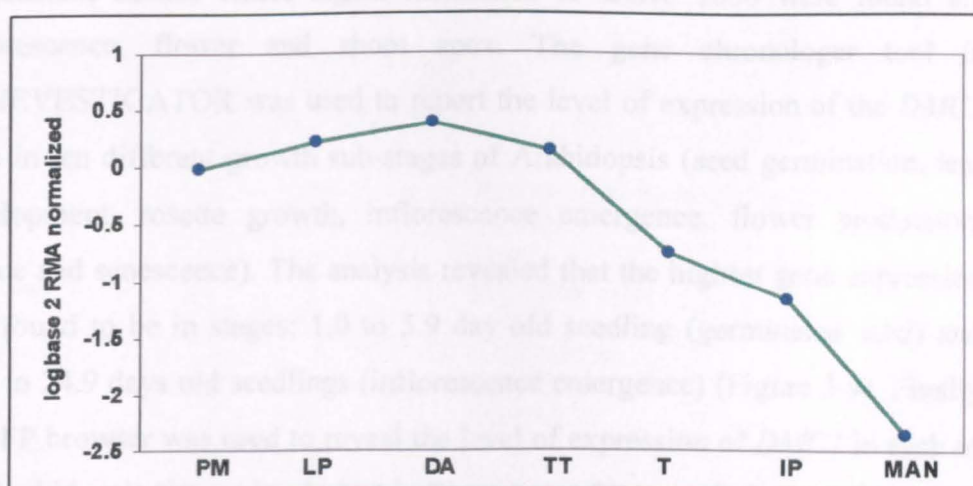
**Figure 3.5** Expression profiling of *OsRAD51A2* gene in different tissues using gene atlas tool in Rice GE database.

The wheat expression dataset for spatial expression of *TaRAD51* and *TaDMC1* genes was obtained from PLEXdb (experiment TA3; Schreiber *et al.*, 2009). The wheat gene atlas tool of PLEXdb revealed that the *TaRAD51* gene is ubiquitously expressed in all developmental tissues starting from germinating seeds to filled grains even though the expression levels vary across the tissues. *TaRAD51* was found to be highly expressed during the immature inflorescence stages with a signal intensity of 11.2 and low expression was found at the anthesis stage with a signal intensity of 2.5 (Figure 3.6). For studying temporal gene expression of *TaRAD51*, the dataset corresponding to the “Microarray expression analysis of meiosis and microsporogenesis in hexaploid wheat” (experiment: GSE6027; Crismani *et al.*, 2006) was only loaded into the Genespring software and not datasets corresponding to TA3 as they were not publicly available. The analysis revealed that *TaRAD51* expression is significantly elevated during the pre-meiosis to diplotene-pachytene of prophase I and there is decrease in *TaRAD51* gene expression levels from diplotene-pachytene to the mature anther stage. As meiosis concludes, transcript levels are minimal (Figure 3.7).





**Figure 3.6** Expression pattern of *TaRAD51* during various stages of wheat development ((experiment TA3; Schreiber *et al.*, 2009) using gene atlas tool in PLEXdb



**Figure 3.7** *TaRAD51* expression levels observed from microarray analysis of meiotic dataset (experiment: GSE6027; Crismani *et al.*, 2006) using Genespring GX10 software. Abbreviations used are PM - Pre-Meiosis; LP - Leptotene-Pachytene; DA - Diplotene-Anaphase I; TT - TelophaseI-TelophaseII; T - Tetrad; IP - Immature Pollen; MAN - Mature Anther.

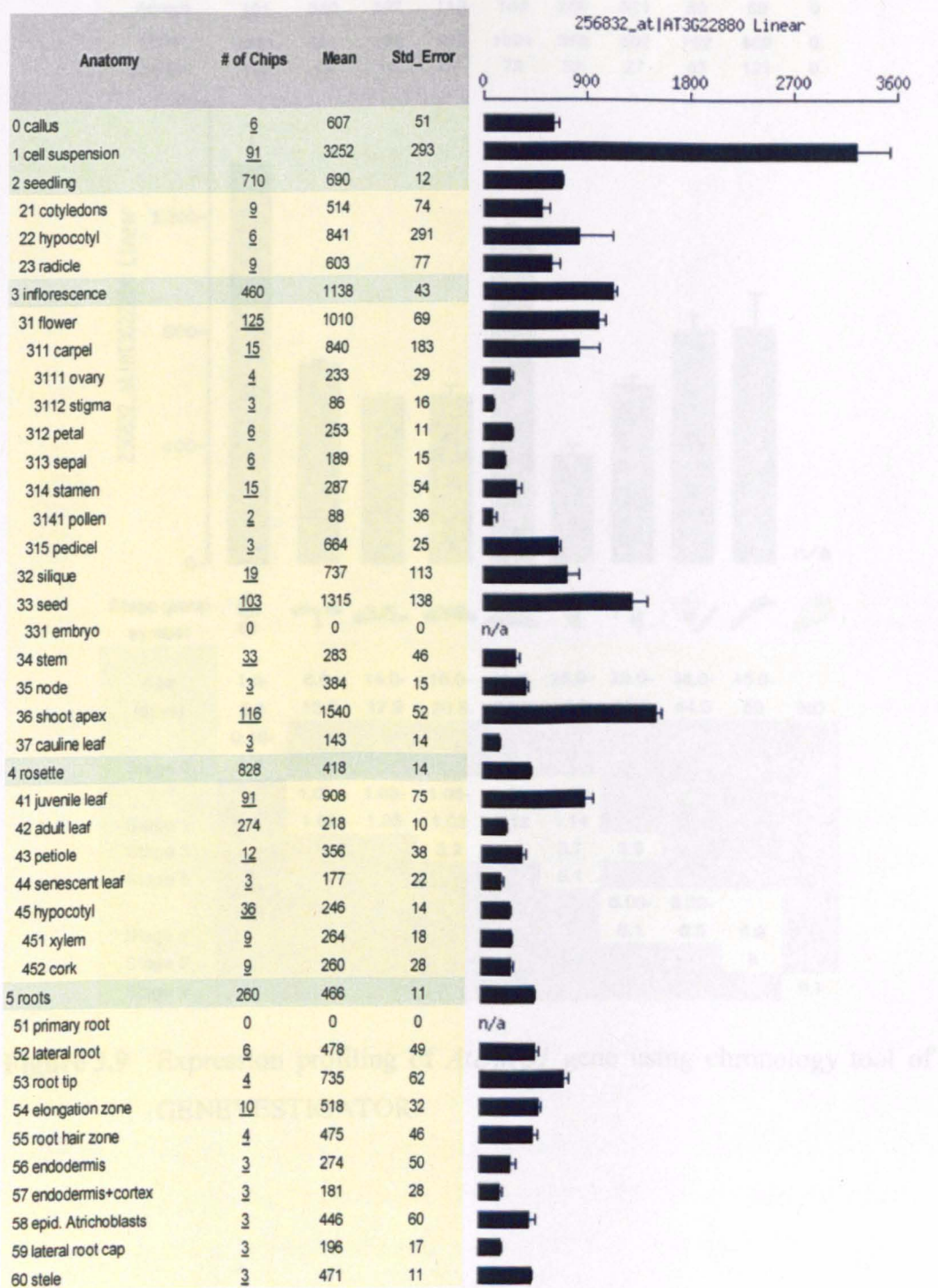
### 3.2.4 Transcriptional expression profiling of the *DMC1* genes in Arabidopsis, rice and wheat using microarray databases

In order to investigate the molecular changes that are brought about by the expression of *DMC1* gene in Arabidopsis, rice and wheat, analysis of gene expression by transcript profiling was done using available expression data in microarray databases such as GENEVESTIGATOR, eFP (for Arabidopsis), RiceGE (for Rice) PLEXdb and Genespring (for wheat).

In Arabidopsis analysis of *DMC1* gene by Gene Atlas tool in GENEVESTIGATOR revealed that the gene is found to be expressed ubiquitously in all developmental tissues even though the levels of expression differed (Figure 3.8). The highest level of *DMC1* gene expression was found to be in cell suspension with a signal intensity value of 3252. Apart from cell

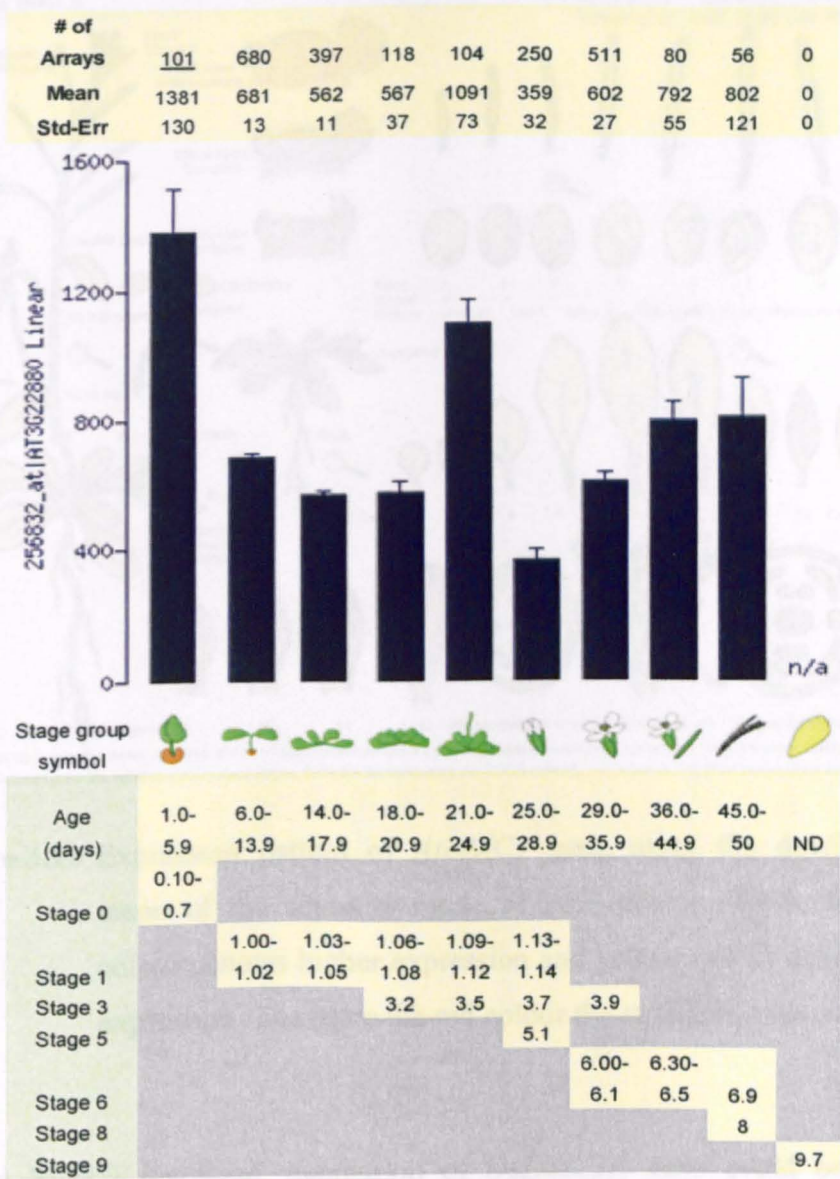
suspension, tissues where signal intensities of above 1000 were found are inflorescence, flower and shoot apex. The gene chronologer tool in GENEVESTIGATOR was used to report the level of expression of the *DMC1* gene in ten different growth sub-stages of Arabidopsis (seed germination, leaf development, rosette growth, inflorescence emergence, flower production, silique and senescence). The analysis revealed that the highest gene expression was found to be in stages: 1.0 to 5.9 day old seedling (germinated seed) and 21.0 to 24.9 days old seedlings (inflorescence emergence) (Figure 3.9). Finally the eFP browser was used to reveal the level of expression of *DMC1* in each of the Arabidopsis tissues in electronic Fluorescent Pictograph form. Arabidopsis eFP browser output for *DMC1* in absolute mode revealed very strong expression in flower (Stage 9) and shoot apex. The lowest level of expression was found to be in leaves, seeds and siliques (Figure 3.10).



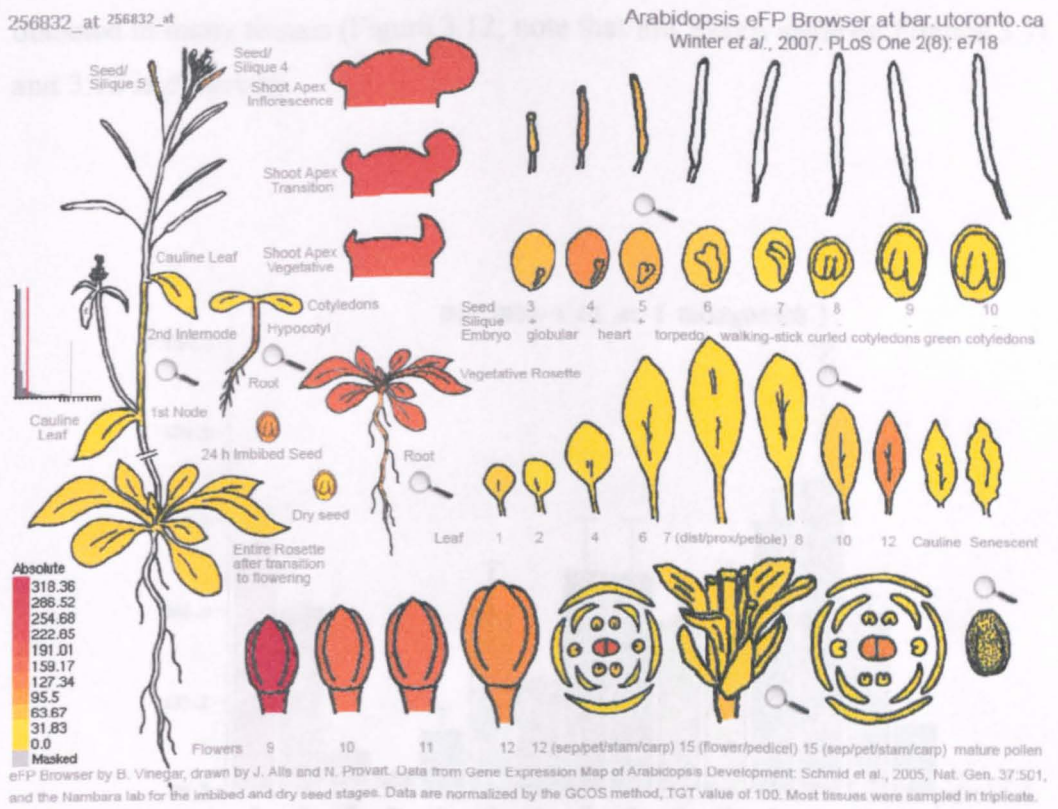


**Figure 3.8** Expression profiling of *AtDMC1* gene using gene atlas tool of GENEVESTIGATOR.





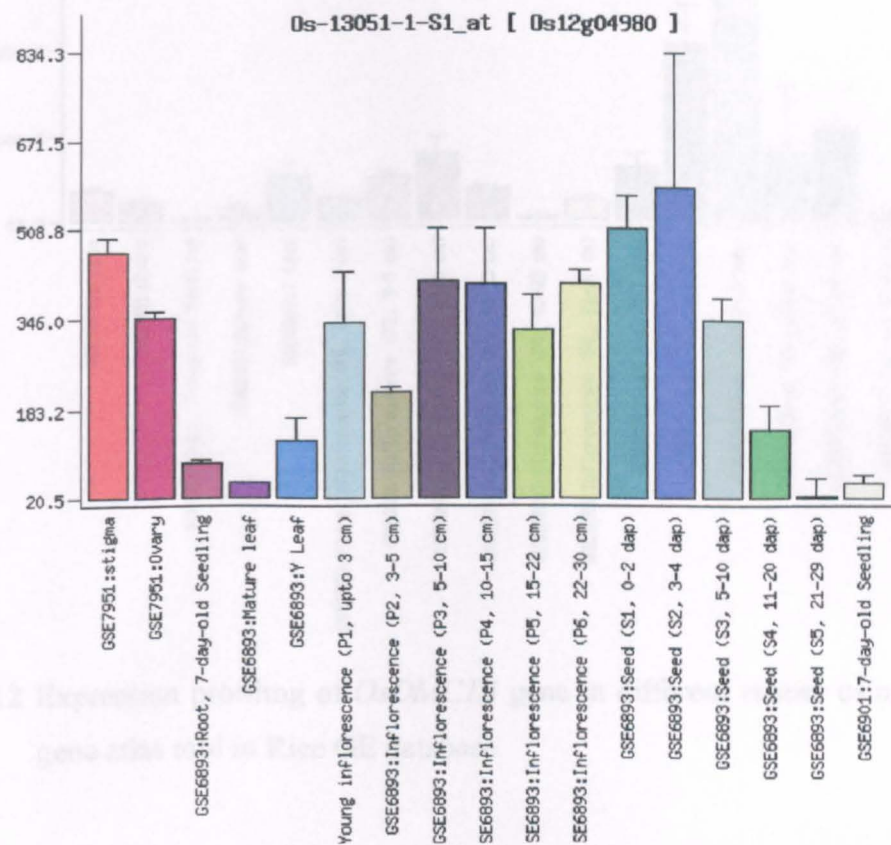
**Figure 3.9** Expression profiling of *AtDMC1* gene using chronology tool of GENEVESTIGATOR.



**Figure 3.10** Expression pattern of *AtDMC1* gene using the developmental menu of the absolute mode of Arabidopsis eFP browser. Red colour denotes higher expression and yellow colour denotes lower expression. The more the red colour the stronger is the expression.

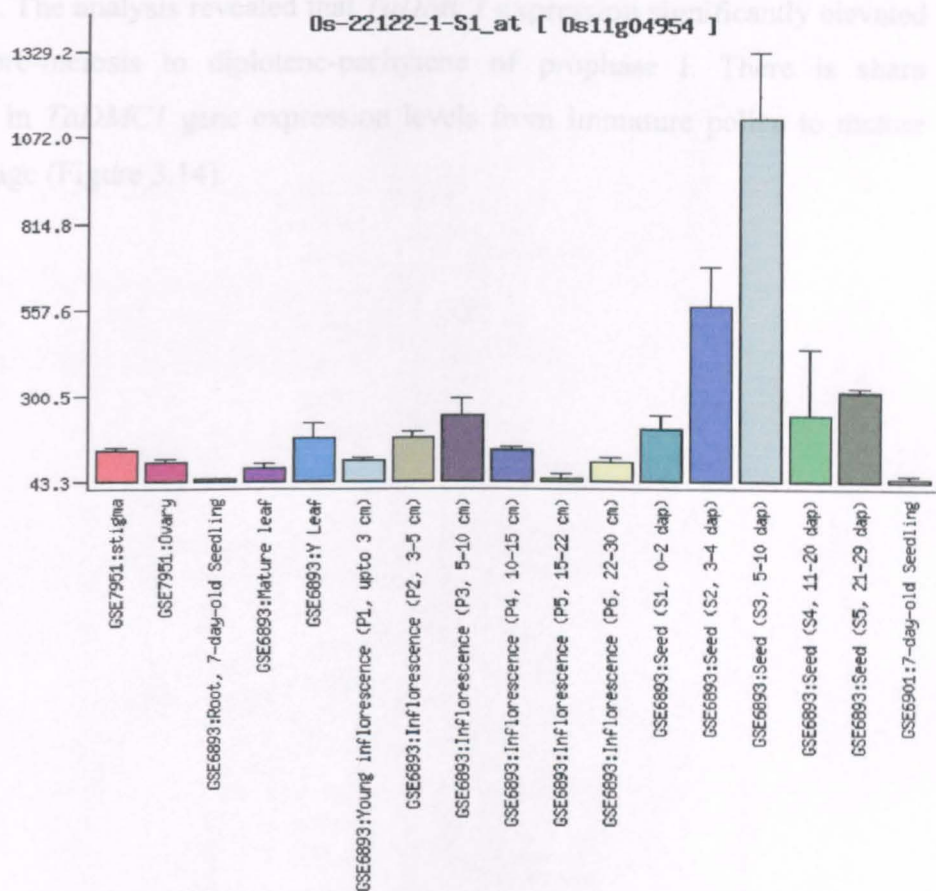
In the RiceGE database, expression of *OsDMC1A* gene could be detected ubiquitously but at variable levels. The higher expression levels of *OsDMC1A* was reported in stigma, inflorescence (5-10 cm, 10-15 cm, 22-30 cm), seed (0-2 dap and 3-4 dap; day after post-anthesis) with the mean signal intensity values from 500 to 590. The lower levels of signal intensity values were reported in 7 day-old seedling, mature leaf, and seed 21-29 dap (Figure 3.11). The highest levels of gene expression of *OsDMC1B* were detected in seed 5-10 dap with a mean of signal of 1,072. In addition, the seed 3-4 dap showed a mean signal of 560, similar to *OsDMC1A*. However, low levels of signal intensity were

detected in many tissues (Figure 3.12; note that the Y-axis scale on Figures 3.11 and 3.12 is different).



**Figure 3.11** Expression profiling of *OsDMC1A* gene in different tissues using gene atlas tool in Rice GE database

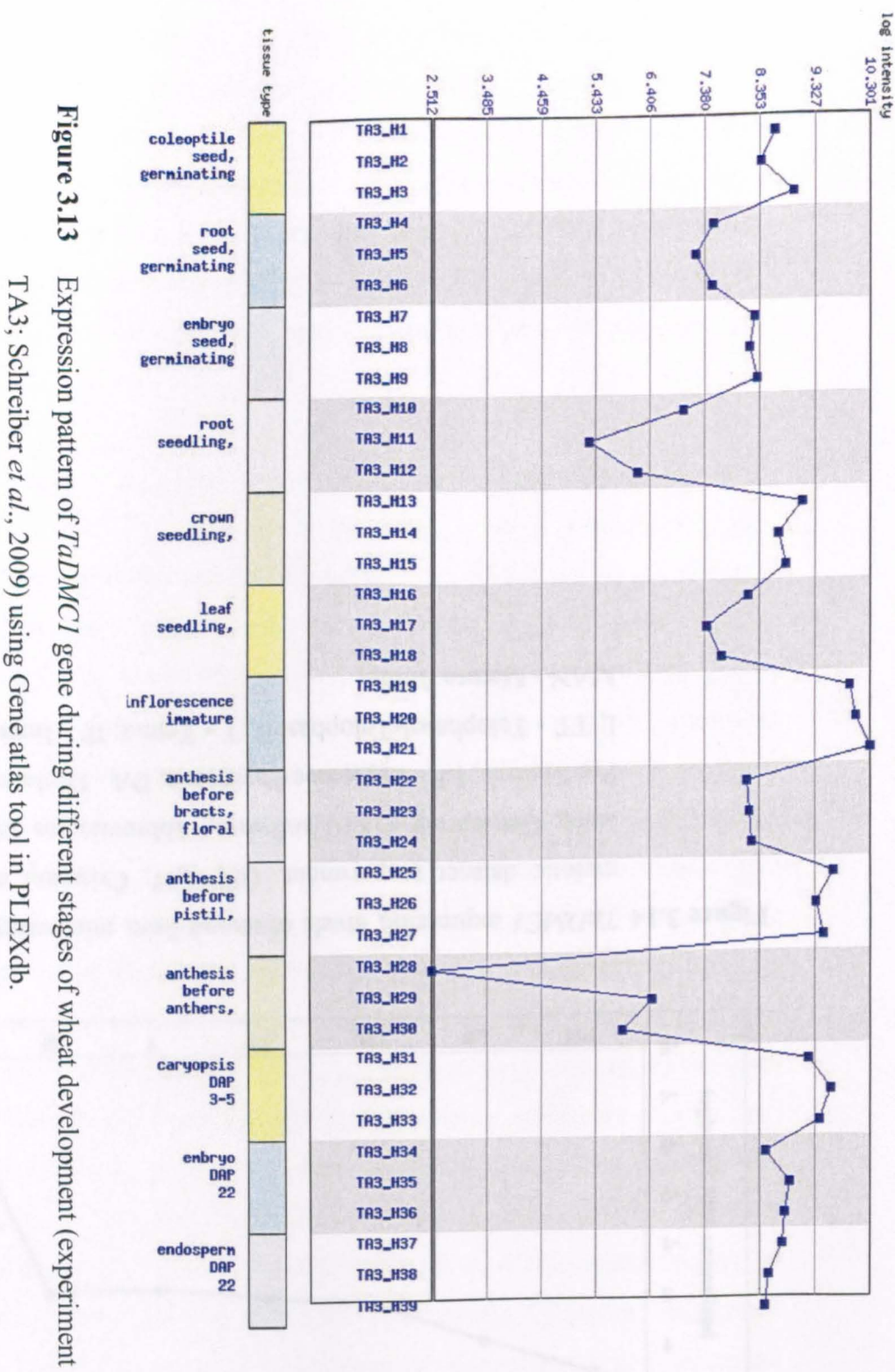


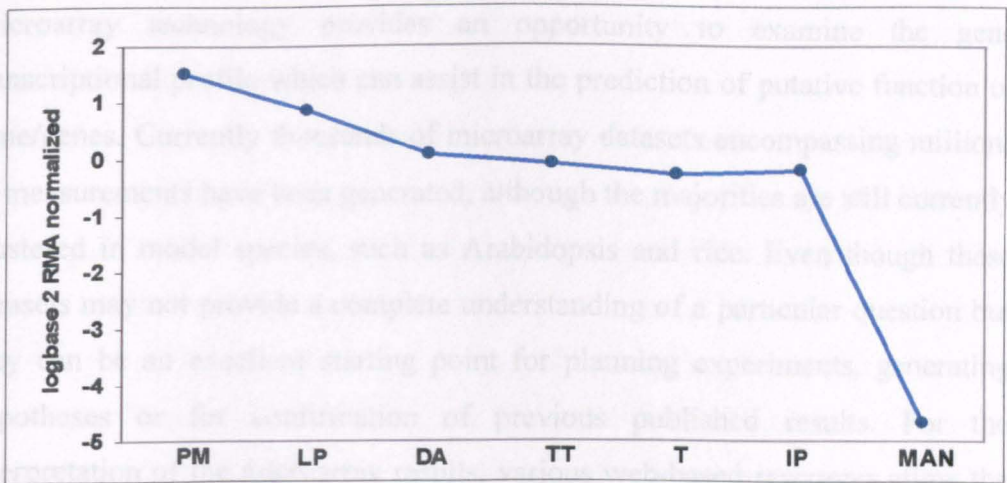


**Figure 3.12** Expression profiling of *OsDMC1B* gene in different tissues using gene atlas tool in Rice GE database.

The wheat gene atlas tool in PLEXdb was used to reveal spatial gene expression of *TaDMC1* in a series of tissues and developmental stages ranging from leaf, crown, caryopsis, anther, pistil, inflorescence, bracts, mesocotyl, endosperm, embryo and coleoptiles. This analysis revealed that the gene is expressed ubiquitously even though the levels of expression varied. *TaDMC1* was also found to be highly expressed during the immature inflorescence stage with a signal intensity of 10.3 and very low expression was found at anthesis stage with a signal intensity of 2.5 (Figure 3.13). For studying temporal gene expression of *TaDMC1*, the dataset corresponding to the “Microarray expression analysis of meiosis and microsporogenesis in hexaploid wheat” (experiment: GSE6027; Crismani *et al.*, 2006) was loaded into Genespring

software. The analysis revealed that *TaDMC1* expression significantly elevated during pre-meiosis to diplotene-pachytene of prophase I. There is sharp decrease in *TaDMC1* gene expression levels from immature pollen to mature anther stage (Figure 3.14).





**Figure 3.14** *TaDMC1* expression levels observed from microarray analysis of meiotic dataset (experiment: GSE6027; Crismani *et al.*, 2006) using Genespring GX10 software. Abbreviations used are PM-Pre-Meiosis; LP - Leptotene-Pachytene; DA - Diplotene-Anaphase I; TT - TelophaseI-TelophaseII; T - Tetrad; IP - Immature Pollen; MAN - Mature Anther.

### 3.3 Discussion

Microarray technology provides an opportunity to examine the gene transcriptional profile which can assist in the prediction of putative function of gene/genes. Currently thousands of microarray datasets encompassing millions of measurements have been generated, although the majorities are still currently clustered in model species, such as *Arabidopsis* and rice. Even though these datasets may not provide a complete understanding of a particular question but they can be an excellent starting point for planning experiments, generating hypotheses or for confirmation of previous published results. For the interpretation of the microarray results, various web-based resources allow the mining of information for each gene analyzed and thereby enabling the comparison of the target sequences of the probe sets with information about the gene structure. Microarray data analyses from this study using online database analysis tools here report the expression profiles of *RAD51* and *DMC1* genes in *Arabidopsis*, rice and wheat. Their expression profiles suggest that their expression is reasonably ubiquitous. This contrasts with the expression of *DMC1* in *Saccharomyces cerevisiae*, where expression is limited to meiotic stages (Shinohara, 2003). The microarray mining data is consistent with RT-PCR and QRT-PCR results in *Arabidopsis*, rice and wheat (Klimyuk and Jones (1997); Doutriaux *et al.*, 1998; Ding *et al.*, 2001; Shimazu *et al.*, (2001); Kathiresan *et al.*, (2002); Osakabe *et al.*, 2002; Khoo *et al.*, 2008).

#### 3.3.1 *RAD51* expression in *Arabidopsis*, rice and wheat

Microarray expression profiling of *AtRAD51* showed expression in both vegetative and reproductive tissues and at many developmental stages. This can be attributed to its role in DNA repair and meiosis and is consistent with earlier reports of studies in *Arabidopsis* using RT-PCR (Doutriaux *et al.*, 1998; Li *et al.*, 2004 and Seo *et al.*, 2007). Moreover, Doutriaux *et al.*, (1998) reported that the strongest level of expression for *AtRAD51* gene appears in RNA extracted from an exponentially growing *Arabidopsis* suspension cultures and microarray



data analysis supported this observation. In RiceGE database, the results of *OsRAD51A1* and *OsRAD51A2* microarray analysis revealed that transcripts are detected in both mitotically active tissue and meiotic tissues. In wheat, *TaRAD51* was found be expressed in both mitotically active tissue and meiotic tissues. Temporal expression analysis patterns revealed that highest expression of *TaRAD51* was confined to the pre-meiosis and diplotene-anaphase. This result is similar to the results reported by Khoo *et al.*, (2008) using a QRT-PCR based approach. They showed that *TaRAD51A1* gene expression was relatively constant at Pre-meiosis and Leptotene-Pachytene but gradually increase from Diplotene-Anaphase I until the Tetrad stage. The highest level of *TaRAD51* gene expression seen during Leptotene to Pachytene and Diplotene-Anaphsel was expected based on these stages being when meiotic recombination is underway.

In summary, the results of *AtRAD51*, *OsRAD151A1*, *OsRAD51A2* and *TaRAD51* expression analysis showed that the *RAD51* gene is expressed in both meiotic and in mitotically-active tissues and suggests it could have a role in mitotic/repair recombination as well as in meiotic exchange.

### **3.3.2 *DMC1* expression in Arabidopsis, rice and Wheat**

For Arabidopsis, a three-fold increase in *DMC1* gene expression was found in cell suspension cultures compared to inflorescence material and high levels of expression were also found in actively dividing tissues such as shoot apex, seed and young leaves. This is similar to the observations by Doutriaux *et al.*, (1998) and Klimyuck and Jones (1997) using an RT-PCR based approach. Based on these two reports, it can be concluded that the microarray expression analysis of *AtDMC1* is consistent with the published results. According to Doutriaux *et al.*, (1998), the increase of *DMC1* expression in suspension cultures could be due to the loss of a repressive promoter element or the inactivation of a regulatory protein that was responsible for non-meiotic repression. Because of the

detection of *AtDMC1* transcripts in root and leaf, it may also suggest that if deregulation is responsible for this widened pattern of expression, it may not simply be restricted to actively dividing cell types found in the suspension cultures. But this needs to be investigated further. An analysis of gene expression profiling of *AtDMC1* during different stages revealed that it was up-regulated through the leaf formation stage to the inflorescence initiation stage. Surprisingly the *AtDMC1* gene showed high levels of expression in the early seedling stage. This could be indicative of a role in seed germination or homologous recombination during mitosis. Microarray expression analysis using RiceGE database suggests that the *OsDMC1A* gene transcript was detected at high levels in young developing seed, medium levels in inflorescence and low levels in seedling tissue. For *OsDMC1B*, the RiceGE database showed high levels of expression after pollination (from R5 to R9 stages). This is similar to the results reported by Kathiresan *et al.*, (2002). Microarray analysis of spatial expression of *TaDMC1* showed that transcripts are detected in both vegetative and reproductive tissues as well. Temporal expression analysis of *TaDMC1* gene clearly indicated expression of *TaDMC1* in meiotic tissue, with the highest expression predominantly confined to pre-meiosis and leptotene to pachytene. The leptotene to pachytene result is consistent with the stages when recombination occurs. This is consistent with the results reported by Khoo *et al.*, (2008) using a QRT-PCR based approach.

In summary, the results of *AtDMC1*, *OsDMC1A*, *OsDMC1B* and *TaDMC1* expression analysis showed that the *DMC1* gene is expressed in both meiotic and mitotically active tissues, especially in flower bud and seedling. This could argue that the *DMC1* gene is not only functional during meiotic homologous recombination but also might be functional in DNA repair by mitotic homologous recombination in a similar fashion to *RAD51*. Both this and the expression of *DMC1* in mature leave and root (non-dividing cell) are intriguing and need to be investigated further.

While microarray expression profiling of *RAD51* and *DMC1* genes for diploid genome such as *Arabidopsis* and *Rice* are straight forward, one complication with microarrays for species such as wheat is the polyploid nature of the genome. For wheat, studying gene expression aimed at individual homoeologues is complicated by genome duplication and the likely high similarity of nucleotide sequences of the three homoeologues, thereby resulting in cross-hybridization of the probes. The reliability of expression profiling of *TaRAD51* and *TaDMC1* genes using the current Affymetrix Wheat chip needs to be interpreted carefully. Depending on the stringency of the probeset design, an expression profile of any gene derived from the Wheat GeneChip may receive contributions from one, two or three homoeologues. The current results suggest that in case of *TaRAD51* gene, out of 11 oligo probes, 8 were found to cross-hybridize to all the three homoeologues and only 3 were found to cross-hybridize 'A' homoeologue. For the *TaDMC1* gene, out of 10 oligo probes, 6 were found to cross-hybridize to all the three homoeologue, 1 probe to both (A) and (B), 1 probe for both (A) and (D) and only 2 probes specify '(D)' homoeologue. So, overall most of the oligos for both *TaRAD51* and *TaDMC1* are found to cross-hybridize to all the three cDNA homoeologues. The detailed study of homoeologue expression patterns across tissues in wheat awaits the construction of microarray platforms that specifically target regions of homoeologue sequence divergence and/or studies employing direct transcriptome sequencing.

In summary, the current study has established that overall there are similarities between the expression patterns of *RAD51* and *DMC1* genes in the three species analyzed (*Arabidopsis*, *rice* and *wheat*). This is partially expected because of their conserved roles during homologous recombination in eukaryotes. Even though microarray experiments provide invaluable information regarding the gene expression pattern and putative functions of known or unknown genes, the results of these datasets need to be considered

carefully and further experimental evidence is needed in order to investigate their putative roles.

## CHAPTER 4: CLONING AND MOLECULAR CHARACTERIZATION OF *RAD51* AND *DMC1* cDNA HOMOELOGUES IN WHEAT (*Triticum aestivum* L.)

### 4.1 Introduction

Meiotic recombination in eukaryotic cells generally requires two homologues of the *E. coli* RecA protein - Rad51 and Dmc1 - which were first discovered in budding yeast (Bishop *et al.*, 1992; Shinohara *et al.*, 1992). Rad51 forms an ATP-dependant nucleofilament with strand-exchange activity, which has been shown to have roles in meiosis, mitosis, homologous recombination and DSB repair (Symington, 2002). Dmc1, a meiotic homologue of Rad51, is important in early recombination events in eukaryotes and is involved in converting double stranded breaks into hybrid joint molecules (Bishop *et al.*, 1992).

*RAD51* and *DMC1* homologues have been identified in diverse organisms such as fungi, animals and plants (Stassen *et al.*, 1997). *In vitro* and *in vivo* studies indicate that these genes are highly conserved across eukaryotes, share common properties and resemble characteristic features of *E. coli* RecA homologue. Very little information is available on the structural and functional properties of the *RAD51* and *DMC1* genes and the derived proteins in plants. The *RAD51* gene exists as a single copy in Arabidopsis, tomato and poppy while maize and *Physcomitrella patens* have two copies each (Stassen *et al.*, 1997; Doutriaux *et al.*, 1998; Franklin *et al.*, 1999; Ayora *et al.*, 2002; Markmann *et al.*, 2002). Bioinformatics searches with the available databases also show the presence of two copies of *RAD51* (*OsRAD51A1* and *OsRAD51A2*) in the genome of rice located on chromosome 11 and 12 respectively. *DMC1* gene sequences have been reported only in a few plant species. While the Arabidopsis and barley genomes contain one copy of the *DMC1* gene (Klimyuck and Jones, 1997),

rice has two copies: *OsDMC1A* and *OsDMC1B*, which are located on chromosomes 12 and 11, respectively (Ding *et al.*, 2001).

Information on the genes and gene homoeologues of *RAD51* and *DMC1* in wheat is very limited and only a single partial cDNA sequence of *DMC1* has been published (Petersen and Seberg, 2002). Recent studies by Khoo *et al.*, (2008) reported the cloning and preliminary molecular characterization of four *RAD51* gene family members in hexaploid wheat (*TaRAD51A1*, *TaRAD51A2*, *TaRAD51C*, *TaRAD51D* and *TaDMC1*). But it was not known whether the homoeologues of these genes were all functional and the level of homology of their coding regions. In particular, the sequences at the N-terminal and C-terminal ends of these proteins are thought to be important for interactions with other HR proteins. The studies undertaken to identify interaction domains within the Rad51 paralogues have shown that any deletion in either the N-terminal or the C-terminal ends of the proteins eliminate protein–protein interactions (Dosanjh *et al.*, 1998; Kurumizaka *et al.*, 2003; Miller *et al.*, 2004) and even a very short deletion can severely disturb the folding of the Rad51 paralogues (Miller *et al.*, 2004).

In this chapter molecular cloning, molecular characterization and functional expression of the *TaRAD51* and *TaDMC1* cDNA homoeologues was undertaken as a first step towards *in planta* functional analyses in hexaploid wheat (*Triticum aestivum* L.). Extensive sequence homology between rice and wheat allowed the cloning of the *TaRAD51* cDNA sequence. Full-length cDNA sequences of the *TaDMC1* cDNA sequence was cloned using primers designed on a published partial *TaDMC1* cDNA sequence. In order to confirm the identity of the *TaRAD51* and *TaDMC1* cDNA sequences, nucleotide sequence alignments, BLASTX searches and phylogenetic analysis were undertaken with other eukaryotic *RAD51* and *DMC1* homologues. The isolation of the cDNA sequences of *TaRAD51* and *TaDMC1* homoeologues was based on sequence differences that existed between individual homoeologues of the respective genes. Finally molecular characterization and functional expression analysis of *TaRAD51* and

*TaDMC1* homoeologues was performed using comparative & structural protein analysis and Q-PCR respectively

.

## 4.2 Results

### 4.2.1 Isolation and molecular cloning of wheat *RAD51* cDNA sequence

The isolation and molecular cloning of the wheat *RAD51* cDNA sequence was performed using 'Highbury' wheat spikelet first strand cDNA. To isolate the wheat *RAD51* cDNA sequence, previously published *OsRAD51A1* cDNA sequence (Gen Bank accession number ABO8O261) was used and a total of four primer sets were designed to amplify different sized fragments covering the full-length cDNA sequence of *OsRAD51A1* (Table 4.1). RT-PCR amplification reactions were performed using Highbury spikelet first strand cDNA with the above primers. Sequencing of the different cloned RT-PCR products and subsequent alignments resulted in the isolation of full-length *RAD51* cDNA sequence of wheat. BLASTX and comparative amino acid analysis showed that the wheat *RAD51* cDNA sequence was highly similar with *OsRAD51A1* and *A2* (90% and 93% respectively) and *AtRAD51* (85%) cDNA sequences. Based on these results, the *RAD51* cDNA sequence was designated as *TaRAD51*. The particular *TaRAD51* cDNA homoeologue (A, B or D) could not be assigned at that time.

**Table 4.1** Primer pairs and their location that were used to clone the wheat *RAD51* cDNA sequence

Region	Sense primer	Antisense primer	Product size (bp)
Exon 1-2	CCTCACATCCCGAGCA TCTC	TCACATCAACTGCAGCTAT TCCAG	171
Exon 8-9	GGTAGTGGCCCAAGTG GATG	TGGTGGTCCAATATCACAT AGGAG	564
Exon 1-7	ATGTCGTCGGCGGCGG	CTCATCCGCTAATTTCTGA AGGC	889
Exon 2-9	AAGTCGATAAGATAAT TGAAGCA	CTGAAACCTTGCTTCAGCT TCAG	879



### **4.2.2 Isolation and molecular cloning of the wheat *DMC1* cDNA sequence**

To isolate the full-length wheat *DMC1* cDNA sequence, RACE was performed on Highbury wheat spikelet first strand cDNA to obtain sequences at the 3' and 5' ends and primers were also designed to the partial *DMC1* cDNA sequence of wheat (Gen bank accession number DQ247845). RT-PCR amplification reactions were performed using Highbury spikelet first strand cDNA with the above primers. Sequencing of the different cloned PCR products and subsequent alignments resulted in the isolation of the full-length wheat *DMC1* cDNA sequence. BLASTX and comparative amino acid analysis showed that the wheat *DMC1* cDNA sequence was highly similar with *OsDMC1A* and *B* (95% and 97%) and *AtDMC1* (81%) cDNA sequences. Based on these results, the *DMC1* cDNA sequence was designated as *TaDMC1*. The particular *TaDMC1* cDNA homoeologue (A, B or D) could not be assigned at that time.

### **4.2.3 Molecular cloning and chromosome location of *TaRAD51* and *TaDMC1* cDNA homoeologues**

#### **4.2.3.1 Development of Genome Specific Primer sets (GSPs) for the *TaRAD51* and *TaDMC1* cDNA sequences**

The Highbury wheat spikelet cDNA amplified using the 5' & 3' UTR primer sets for *TaRAD51* (Table 4.2) produced a fragment of size of about 1500bp. In the absence of different sized fragments using the full-length primers, the single fragment was then gel eluted, cloned and 15 clones, expected to contain a mixture of homoeologues, with correct insert size as determined by colony PCR were sequenced. Only one sized fragment was obtained in all the clones as expected but sequence analysis of the all the 15 clones identified 3 distinct sequences that differed mostly at 3'UTR sequence sites for the *TaRAD51* cDNA sequence. Putative sets of Genome-Specific primers were designed based on the sequence differences identified and coupled with a generic primer (where there is no sequence difference between homoeologues) (Figure 4.1 and Table 4.2). Nulli-Tetrasomic analysis with the Genome-Specific primers for *TaRAD51* revealed that the

*TaRAD51* cDNA sequence was located on Group 7 chromosomes with a copy on each of the A, B and D genomes of hexaploid wheat (Figure 4.3).

The Highbury wheat spikelet cDNA amplified using the 5' & 3' UTR primer sets (Table 4.2) produced a fragment of size of 1500bp for *TaDMC1* cDNA sequence. In the absence of different sized fragments using the full-length primers, the single fragment were then gel eluted, cloned and 18 clones with correct insert size as determined by colony PCR were sequenced. Only one sized fragment was obtained in all the clones as expected but the sequence analysis of the all the 18 clones identified 3 distinct sequences that differed mostly at 5'UTR sequence sites for the *TaDMC1* cDNA sequence. Putative sets of Genome-Specific primers were designed based on sequence differences identified and coupled with a generic primer (where there is no sequence difference between homoeologues) (Figure 4.2 and Table 4.2). Nulli-Tetrasomic analysis with the Genome-Specific primers for *TaRAD51* revealed that the *TaRAD51* cDNA sequence was located on Group 5 chromosomes with a copy on each of the A, B and D genomes of hexaploid wheat (Figure 4.4).



```

AGCCTCCGCCCCACTTCCTTCTCCTCCAGCAGCACGCGATCCTCCGGCTCGTACAGGTCT 60
AGCCTCCGCCCCACTTC-----TCCTCCAGCAGCACGCGAACCTCCGCCTCTTACAGGTCT 56
AGCCTTGCCCCACTTCCT-CTCCTCTAGCAGCACACGATCCTCTGCCTCGTACAGGTCT 59
*****
CGTCCTCCTCCACCGCTTC CACTC GCCGTCTTCGCTCGTGC GTCC CAGTGCTC GCTC CTT 120
CGTCCTCCTCCACCGCTTC CACTC GCCGTCTTCGCTCGCGTGTCC CAGTGCTC GCTC CTG 116
CGTCCTCCTCCACCGCTTC CACTC GCCGTCTTCGCTCGTGC GTCC CAGAGCTC GCTC CTG 119
*****

ACACCTTCTCCTGTGACGCAGGCGCGCACGCGGGATGGCGC CGTC CAAGCAGTACGACGA 180
ACACCTTCTCCTGTGACGCAGGCGGACACGCGGGATGGCGC CGTC CAAGCAGTACGACGA 176
ACACCTTTTCTCCTGTGACGCAGGCGGGCACGCGGGATGGCGC CGTC CAAGCAGTACGACGA 179
*****

GGGCGGGCAGCTCCAGCTCATGGAGGCCGACC GGGTCGAGGAGGAGGAGGAGTGCTTCGA 240
GGGCGGGCAGCTCCAGCTCATGGAGGCCGACC GGGTCGAGGAGGAGGAGGAGTGCTTCGA 236
GGGCGGGCAGCTCCAGCTCATGGAGGCCGACC GGGTCGAGGAGGAGGAGGAGTGCTTCGA 239
*****

GTCCATCGACAAG 253
GTCCATCGACAAG 249
GTCCATCGACAAG 252
*****

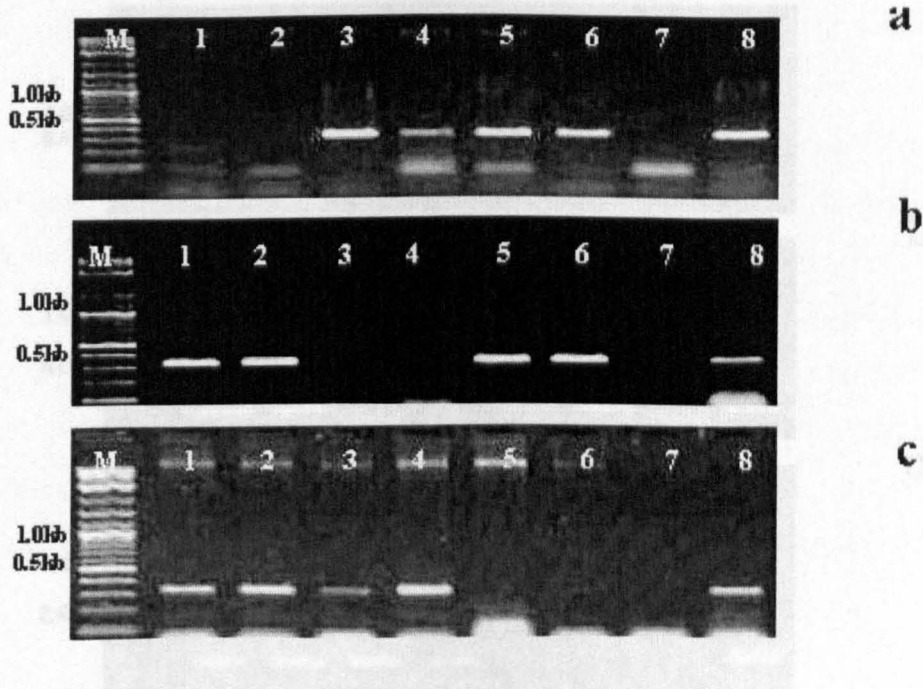
```

**Figure 4.2** Alignment of the 5' upstream region of three *TaDMC1* cDNA homoeologues and design of Genome-Specific primers. The three different coloured letters (Red, blue and green) represent the three putative genome-specific primers and bold letters represent the generic primer. The red arrows indicate the position of forward and reverse primers. Only part of the sequence is shown here

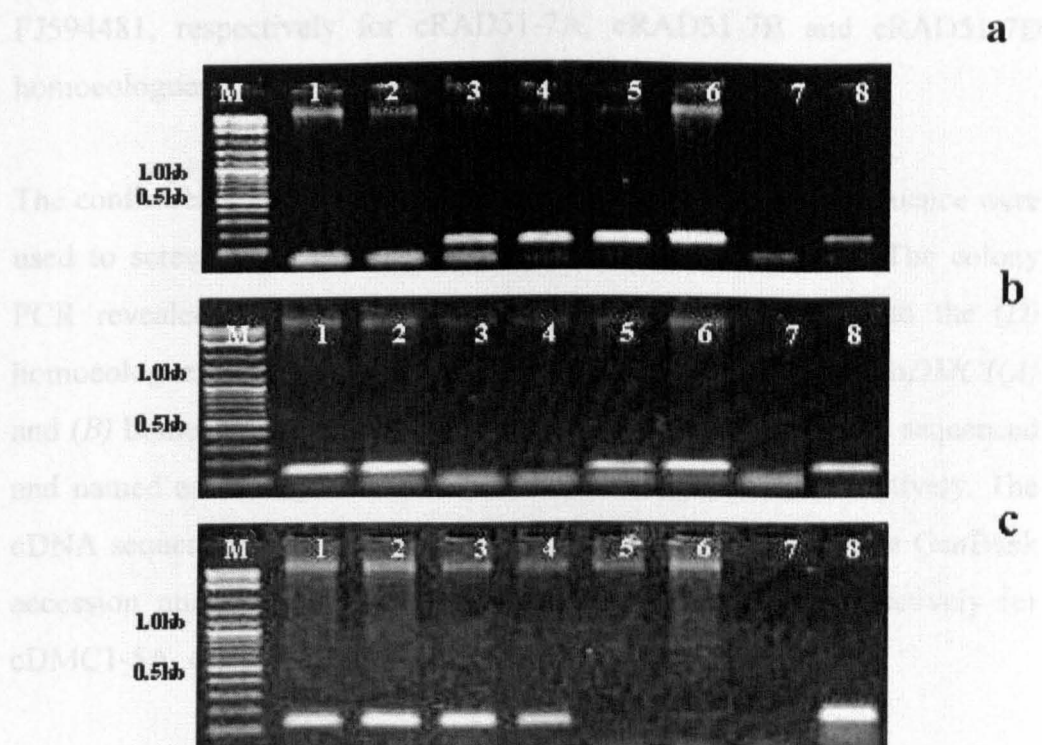
**Table 4.2** UTR and Genome-Specific primers (GSP) used for amplification and sequencing of *TaRAD51* and *TaDMC1* cDNA homoeologues. The *TaRAD51* and *TaDMC1* GSP primers were also used to determine expression profiles in mitotic and meiotic tissues.

Gene name	Sequence (5'-3') Forward	Reverse	T <sub>m</sub> (°C)	Product length (bp)
<b><i>RAD51</i></b>				
UTR primers	CCTCACATCCCGAGCAT CTC	CTCCTATGTGATATTGGACCA CCA	63	1528
A-specific primers	GGGGATACCTCGTGTA TCAGACTT	TGGTGGTCCAATATCACATA GGAG	58	328
B-specific primers	GATACATCGTGTATCG GACA	TGGTGGTCCAATATCACATA GGAG	58	328
D-specific primers	TGGGGATACATCGTGT ATTGGCCT	TGGTGGTCCAATATCACATA GGAG	65	327
<b><i>DMC1</i></b>				
UTR primers	ATGATCCACATTCCACC CGC	CACTGCAGAAAAGAAATTGG GCAAC	63	1268
A-specific primers	AGCCTCCGCCCCACTTC CTTC	CTTGTCGATGGACTCGAAGC ACTC	65	150
B-specific primers	AGCCTTGCCCCACTTC CTC	CTTGTCGATGGACTCGAAGC ACTC	65	150
D-specific primers	TTCTCCTCCAGCAGCAC GCGAA	CTTGTCGATGGACTCGAAGC ACTC	65	150
<b>Q-PCR Control primers</b>				
Tubulin primers	TCTTCATGGTGGGCTTC GC	CGCCTCGGGTGA ACTCCATCT	55	475





**Figure 4.3** PCR assay for the confirmation of *TaRAD51* Genome-Specific primers with Nulli-tetrasomic lines for Group 7 homoeologues. *TaRAD51(A)* Genome -Specific primers (a), *TaRAD51(B)* genome specific primers (b) and *TaRAD51(D)* genome specific primers (c); 1-NT7A7B, 2-NT7A7D, 3-NT7B7A, 4-NT7B7D, 5-NT7D7A, 6-NT7D7B, 7-Water control and 8-Highbury control; M, 2-log ladder.



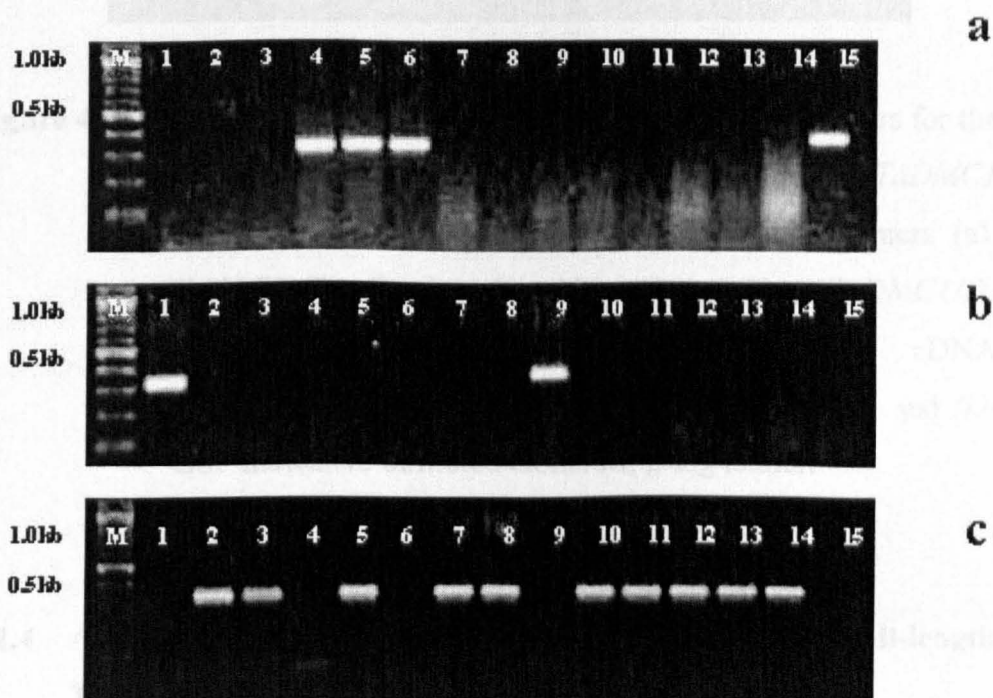
**Figure 4.4** PCR assay for the confirmation of *TaDMC1* Genome-Specific primers with Nulli-tetrasomic lines for Group 5 homoeologues. *TaDMC1(A)* Genome-Specific primers (a), *TaDMC1(B)* Genome-Specific primers (b) and *TaDMC1(D)* Genome-Specific primers (c); 1-NT5A5B, 2-NT5A5D, 3-NT5B5A, 4-NT5B5D, 5-NT5D5A, 6-NT5D5B, 7-Water control and 8-Highbury control; M, 2-log ladder.

#### 4.2.3.2 Isolation of the full-length *TaRAD51* and *TaDMC1* cDNA homoeologues

The confirmed Genome-Specific primers of *TaRAD51* cDNA sequence were used to screen the 15 clones of *TaRAD51* cDNA sequence previously cloned. Colony PCR showed that a majority of the clones (10) belonged to the (D) homoeologue of *TaRAD51* and only 3 and 2 belonged to *TaRAD51(A)* and (B) homoeologue, respectively (Figure 4.5). The clones were sequenced and named as cRAD51-7A, cRAD51-7B and cRAD51-7D respectively. The cDNA sequences were submitted to NCBI and were assigned the GenBank accession numbers: FJ594479, FJ594480 and

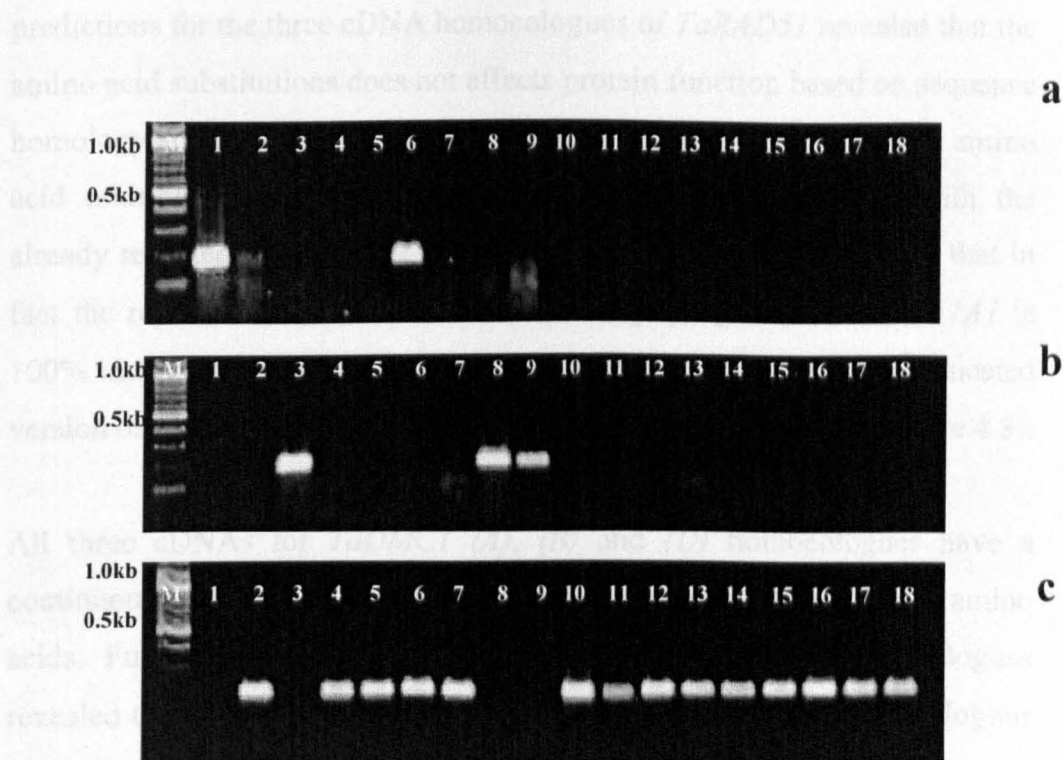
FJ594481, respectively for cRAD51-7A, cRAD51-7B and cRAD51-7D homoeologues.

The confirmed Genome-Specific primers of *TaDMC1* cDNA sequence were used to screen the 18 clones of *TaDMC1* previously cloned. The colony PCR revealed that a majority of the clones (14) belonged to the (*D*) homoeologue of *TaDMC1* and only 1 and 3 clones belonged to *TaDMC1(A)* and (*B*) homoeologue respectively (Figure 4.6). The clones were sequenced and named as cDMC1-5A, cDMC1-5B and cDMC1-5D, respectively. The cDNA sequences were submitted to NCBI and were assigned the GenBank accession numbers: FJ594476, FJ594477 and FJ594478, respectively for cDMC1-5A, cDMC1-5B and cDMC1-5D.



**Figure 4.5** Colony PCR with *TaRAD51* Genome-Specific primers for the isolation of full-length cDNA sequences of the *TaRAD51* cDNA homoeologues. *TaRAD51(A)* Genome-Specific primers (a), *TaRAD51(B)* Genome-Specific primers (b), *TaRAD51(D)* Genome-Specific primers (c); 1-15: *TaRAD51* gene clones; Colony 5 is amplified by both (*A*) and (*D*) GSP indicative of mixed clone; M, 2-log ladder.





**Figure 4.6** Colony PCR with *TaDMC1* Genome-Specific primers for the isolation of full-length cDNA sequences of the *TaDMC1* homoeologues. *TaDMC1(A)* Genome-Specific primers (a), *TaDMC1(B)* Genome-Specific primers (b), *TaDMC1(D)* Genome-Specific primers (c); 1-18:*TaDMC1* cDNA sequence clones; Colony 6 is amplified by both (A) and (D) GSP indicative of mixed clone; M, 2-log ladder.

#### 4.2.4 Analysis of amino acid sequence homology of full-length *TaRAD51* and *TaDMC1* cDNA homoeologues

Two of the *TaRAD51* cDNA clones (*TaRAD51(A)* and (*B*)) have a continuous ORF of 1032bp, capable of encoding a protein of 343 amino acids and the other cDNA clone *TaRAD51(D)* has a deletion of three nucleotides at the N-terminal end thus putatively encoding a predicted protein of 342 amino acids residues. Further analysis of the three translated *TaRAD51* homoeologues revealed that there is a deletion of one amino acid at position 17 corresponding to an 'E' residue (Glutamic acid) and single amino acid substitutions at positions 4, 31 and 115 (Figure 4.7). SIFT

predictions for the three cDNA homoeologues of *TaRAD51* revealed that the amino acid substitutions does not affects protein function based on sequence homology and the physical properties of amino acids (Table 4.3). The amino acid multiple alignment of the three *TaRAD51* homoeologues with the already reported *TaRAD51A1* and *TaRAD51A2* amino acid revealed that in fact the reported cDNA sequences were not paralogues as *TaRAD51A1* is 100% identical to *TaRAD51(D)* homoeologue and *TaRAD51A2* is truncated version of *TaRAD51(A)* albeit with few differences in sequence (Figure 4.8).

All three cDNAs for *TaDMC1* (*A*), (*B*) and (*D*) homoeologues have a continuous ORF of 1035 bp which encodes a predicted protein of 344 amino acids. Further analysis of the three translated *TaDMC1* homoeologues revealed that the differences found between the three cDNA homoeologous sequences were due to single amino acid substitutions at positions 114, 166, 214, 310 and 316; no insertions or deletions were detected (Figure 4.9). Multiple alignments of the amino acid sequence of the three cDNA homoeologues of *TaDMC1* cloned in this thesis with *TaDMC1* reported by Khoo *et al.*, (2008) revealed that the *TaDMC1* is the (*D*) homoeologue (Figure 4.10). SIFT predictions for the three cDNA homoeologues of *TaDMC1* revealed that the amino acid substitutions does not affects protein function based on sequence homology and the physical properties of amino acids (Table 4.4).

Protein alignment was performed to compare the deduced amino acid sequences of *TaRAD51* and *TaDMC1* cDNA homoeologues. Sequence analysis of *TaRAD51* cDNA homoeologues revealed that there is a high level of conservation at the amino acid level. Between individual cDNA homoeologues, the level of conservation is 99%. For *TaDMC1* the level of conserved amino acids is 98%. This suggests functional redundancy of these cDNA homoeologues in the wheat genome is likely, if all are expressed.





TaRAD51A1	MSSSAAAHQKAAAAAPV-EEEEAGEHGPFPIEHLQASGIAAVDVKKLKDAGLCTVESVAYSP	59
TaRAD51D	MSSSAAAHQKAAAAAPV-EEEEAGEHGPFPIEHLQASGIAAVDVKKLKDAGLCTVESVAYSP	59
TaRAD51B	MSSSAAAHQKAAAAAPVEEEAGEHGPFPIEHLQASGIAAVDVKKLKDAGLCTVESVAYSP	60
TaRAD51A	MSSSAAAHQKAAAAAPVEEEAGEHGPFPIEHLQASGIAAVDVKKLKDAGLCTVESVAYSP	60
TaRAD51A2	MSSSAAAHQKAAAAAPVEEEAGEHGPFPIEHLQASGIAAVDVKKLKDAGLCTVESVAYSP	60
*****:*****:*****		
TaRAD51A1	RKDLLQIKGISEAKVDKII EAASKLVPLGFTSATQLHAQRLEIIQVTTGSRELDKILEGG	119
TaRAD51D	RKDLLQIKGISEAKVDKII EAASKLVPLGFTSATQLHAQRLEIIQVTTGSRELDKILEGG	119
TaRAD51B	RKDLLQIKGISEAKVDKII EAASKLVPLGFTSATQLHAQRLEIIQVTTGSRELDRILEGG	120
TaRAD51A	RKDLLQIKGISEAKVDKII EAASKLVPLGFTSATQLHAQRLEIIQVTTGSRELDKILEGG	120
TaRAD51A2	RKDLLQIKGISEAKVDKII EAASKLVPLGFTSATQLHAQRLEIIQVTTGSRELDKILEGG	120
*****:*****		
TaRAD51A1	IETGSITELYGEFRSGKTQLCHTLCVTCQLPLDQGGGEGKALYIDAEGTFRPQRLQIAD	179
TaRAD51D	IETGSITELYGEFRSGKTQLCHTLCVTCQLPLDQGGGEGKALYIDAEGTFRPQRLQIAD	179
TaRAD51B	IETGSITELYGEFRSGKTQLCHTLCVTCQLPLDQGGGEGKALYIDAEGTFRPQRLQIAD	180
TaRAD51A	IETGSITELYGEFRSGKTQLCHTLCVTCQLPLDQGGGEGKALYIDAEGTFRPQRLQIAD	180
TaRAD51A2	IETGSITELYGEFRSGKTQLCHTLCVTCQLPLDQGGGEGKALYIDAEGTFRPQRLQIAD	180
*****		
TaRAD51A1	RFGLNGADVLENVAYARAYNTDHQSRLLEAASMMVETRFALMVIDSATALYRTDFSGRG	239
TaRAD51D	RFGLNGADVLENVAYARAYNTDHQSRLLEAASMMVETRFALMVIDSATALYRTDFSGRG	239
TaRAD51B	RFGLNGADVLENVAYARAYNTDHQSRLLEAASMMVETRFALMVIDSATALYRTDFSGRG	240
TaRAD51A	RFGLNGADVLENVAYARAYNTDHQSRLLEAASMMVETRFALMVIDSATALYRTDFSGRG	240
TaRAD51A2	RFGLYADVLENVAYARAYNTDHQSRLLEAASMMVETRFALMVIDSATALYRTDFSGRG	240
**** *		
TaRAD51A1	EL SARQMHLAKFLRSLQKLAD EFGVAVVISNQVVAQVDGGAMFAGPQIKPIGGNIMAHAS	299
TaRAD51D	EL SARQMHLAKFLRSLQKLAD EFGVAVVISNQVVAQVDGGAMFAGPQIKPIGGNIMAHAS	299
TaRAD51B	EL SARQMHLAKFLRSLQKLAD EFGVAVVISNQVVAQVDGGAMFAGPQIKPIGGNIMAHAS	300
TaRAD51A	EL SARQMHLAKFLRSLQKLAD EFGVAVVISNQVVAQVDGGAMFAGPQIKPIGGNIMAHAS	300
TaRAD51A2	EL SARQMHLAKFLRSLQKLAD EFGVAVVISNQVVAQVDGGAMFAGPQIKPIGGNIMAHAS	300
*****		
TaRAD51A1	TTRLYLRLKGRAEERICKVVSSPCLAEAEARFQISPEGVTDVKD	342
TaRAD51D	TTRLYLRLKGRAEERICKVVSSPCLAEAEARFQISPEGVTDVKD	342
TaRAD51B	TTRLYLRLKGRAEERICKVVSSPCLAEAEARFQISPEGVTDVKD	343
TaRAD51A	TTRLYLRLKGRAEERICKVVSSPCLAEAEARFQISPEGVTDVKD	343
TaRAD51A2	TTRLYLRLKGKGGGADL-----	316
*****:* :*		

**Figure 4.8** Alignment and comparison of the deduced amino acids of the *TaRAD51* cDNA homoeologues with already reported *TaRAD51A1* and *A2* paralogues. Conserved amino acids are indicated by black with a yellow background. The amino acid similarities between the *TaRAD51A1* and *TaRAD51(D)* is indicated by black with green background and similarities between *TaRAD51A2* and *TaRAD51(A)* is indicated by black with grey background.

```

TaDMC1A  MAPSKQYDEGGQLQLMEADRVVEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMM 60
TaDMC1D  MAPSKQYDEGGQLQLMEADRVVEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMM 60
TaDMC1B  MAPSKQYDEGGQLQLMEADRVVEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMM 60
*****

TaDMC1A  HTKKSILTGIKGLSEAKVDKICEAAEKLLSQGFMTGSDLLIKRKSVVRIITTCGQALDELLG 120
TaDMC1D  HTKKSILTGIKGLSEAKVDKICEAAEKLLSQGFMTGSDLLIKRKSVVRIITTCGQALDELLG 120
TaDMC1B  HTKKSILTGIKGLSEAKVDKICEAAEKLLSQGFMTGSDLLIKRKSVVRIITTCGQALDELLG 120
*****

TaDMC1A  GGIETLCITEAFGEFRSGKTQLAHTLCVSTQLPLMHGCGNCKVAYIDTEGTFRPERIVPI 180
TaDMC1D  GGIETLCITEAFGEFRSGKTQLAHTLCVSTQLPLMHGCGNCKVAYIDTEGTFRPERIVPI 180
TaDMC1B  GGIETLCITEAFGEFRSGKTQLAHTLCVSTQLPLMHGCGNCKVAYICTEGTFRPERIVPI 180
*****

TaDMC1A  AERFCMDANAVLDNIIYARAYTYEHQYNLLCLAAKMAEPPRLLIIVDSVIALFRVDFSG 240
TaDMC1D  AERFCMDANAVLDNIIYARAYTYEHQYNLLCLAAKMAEPPRLLIIVDSVIALFRVDFSG 240
TaDMC1B  AERFCMDANAVLDNIIYARAYTYEHQYNLLCLVAKMAEPPRLLIIVDSVIALFRVDFSG 240
*****

TaDMC1A  RCELAERQQKLAQMLSRLTKIAEEFNVAUYITNQVIADPGCGMFIIDPKKPAGCHVLAHA 300
TaDMC1D  RCELAERQQKLAQMLSRLTKIAEEFNVAUYITNQVIADPGCGMFIIDPKKPAGCHVLAHA 300
TaDMC1B  RCELAERQQKLAQMLSRLTKIAEEFNVAUYITNQVIADPGCGMFIIDPKKPAGCHVLAHA 300
*****

TaDMC1A  ATIRLMRLRKSGEQRVCKIFDAPNLPEGEAVFQITTCGLMDVKD 344
TaDMC1D  ATIRLMRLRKSGEQRVCKIFDAPNLPEGEAVFQITTCGLMDVKD 344
TaDMC1B  ATIRLMRLRKSGEQRVCKIFDAPNLPEGEAVFQITTCGLMDVKD 344
*****

```

**Figure 4.9** Alignment and comparison of the deduced amino acids of the *TaDMC1* cDNA homoeologues. Conserved amino acids are indicated by black with a yellow background. The amino acid differences between the three cDNA homoeologues are indicated by a black with grey background.

**Table 4.4** SIFT predictions for the amino acid substitutions for the three cDNA homoeologues of *TaDMC1*

SNP Position	<i>TaDMC1</i> homoeologues			SIFT predictions
	A	B	D	
114	A	T	A	Tolerated
167	D	D	G	Tolerated
214	A	A	V	Tolerated
310	S	G	G	Tolerated
316	V	V	I	Tolerated



TaDMC1-5D	MAPSKQYDEGGQLQLMEADRVEEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMM	60
TaDMC1	MAPSKQYDEGGQLQLMEADRVEEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMM	60
TaDMC1-5A	MAPSKQYDEGGQLQLMEADRVEEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMM	60
TaDMC1-5B	MAPSKQYDEGGQLQLMEADRVEEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMM	60
*****		
TaDMC1-5D	HTKKSALTGIKGLSEAKVDKICEAAEKLLSQGFMTGSDLLIKRKSVMRITTSQTLDELLE	120
TaDMC1	HTKKSALTGIKGLSEAKVDKICEAAEKLLSQGFMTGSDLLIKRKSVMRITTSQTLDELLE	120
TaDMC1-5A	HTKKSALTGIKGLSEAKVDKICEAAEKLLSQGFMTGSDLLIKRKSVMRITTSQALDELLE	120
TaDMC1-5B	HTKKSALTGIKGLSEAKVDKICEAAEKLLSQGFMTGSDLLIKRKSVMRITTSQALDELLE	120
*****		
TaDMC1-5D	GGIETLCITEAFGEFRSGKTQLAHTLCVSTQLPLMHGGNGKVAYIDTEGTFRPERIVPI	180
TaDMC1	GGIETLCITEAFGEFRSGKTQLAHTLCVSTQLPLMHGGNGKVAYIDTEGTFRPERIVPI	180
TaDMC1-5A	GGIETLCITEAFGEFRSGKTQLAHTLCVSTQLPLMHGGNGKVAYIDTEGTFRPERIVPI	180
TaDMC1-5B	GGIETLCITEAFGEFRSGKTQLAHTLCVSTQLPLMHGGNGKVAYIGTEGTFRPERIVPI	180
*****		
TaDMC1-5D	AERFGMDANAVLDNIIYARAYTYEHQYNLLGLAAKMAEFPRLIVDSVIALFRVDFSG	240
TaDMC1	AERFGMDANAVLDNIIYARAYTYEHQYNLLGLAAKMAEFPRLIVDSVIALFRVDFSG	240
TaDMC1-5A	AERFGMDANAVLDNIIYARAYTYEHQYNLLGLAAKMAEFPRLIVDSVIALFRVDFSG	240
TaDMC1-5B	AERFGMDANAVLDNIIYARAYTYEHQYNLLGLVAKMAEFPRLIVDSVIALFRVDFSG	240
*****		
TaDMC1-5D	RGELAERQQKLAQMLSRITKIAEEFNNAVYITNQVIADPGGGMFITDPKKPAGGHVLAHA	300
TaDMC1	RGELAERQQKLAQMLSRITKIAEEFNNAVYITNQVIADPGGGMFITDPKKPAGGHVLAHA	300
TaDMC1-5A	RGELAERQQKLAQMLSRITKIAEEFNNAVYITNQVIADPGGGMFITDPKKPAGGHVLAHA	300
TaDMC1-5B	RGELAERQQKLAQMLSRITKIAEEFNNAVYITNQVIADPGGGMFITDPKKPAGGHVLAHA	300
*****		
TaDMC1-5D	ATIRLMRLRKKGQQRVCKIFDAPNLPEGEAVFQITTGGLMDVKD	344
TaDMC1	ATIRLMRLRKKGQQRVCKIFDAPNLPEGEAVFQITTGGLMDVKD	344
TaDMC1-5A	ATIRLMRLRKSGQQRVCKIFDAPNLPEGEAVFQITTGGLMDVKD	344
TaDMC1-5B	ATIRLMRLRKKGQQRICKIFDAPNLPEGEAVFQITTGGLMDVKD	344
*****		

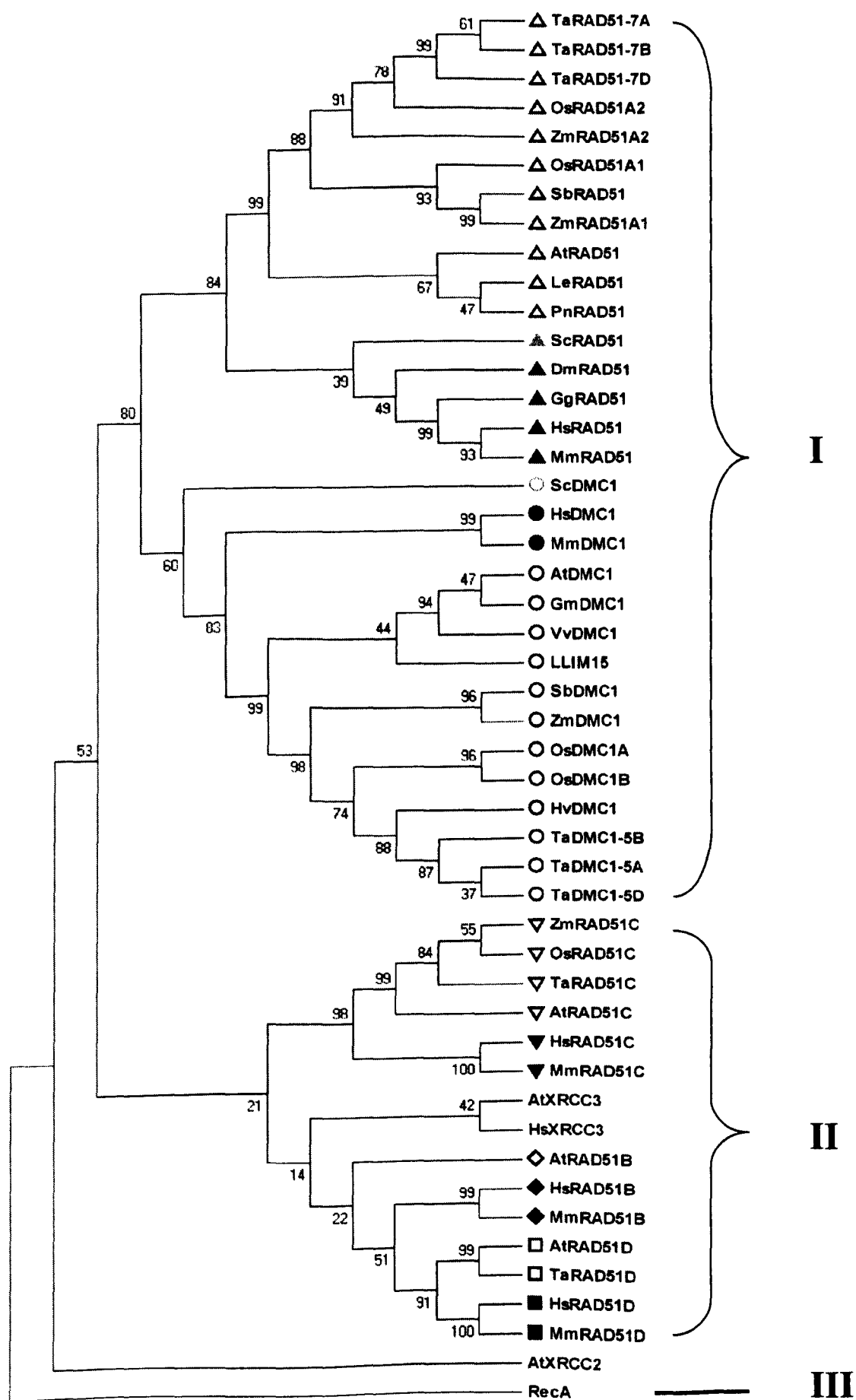
**Figure 4.10** Alignment and comparison of the deduced amino acids of the *TaDMC1* cDNA homoeologues with already reported *TaDMC1*. Conserved amino acids are indicated by black with a yellow background. The amino acid similarities between the *TaDMC1* and *TaDMC1(D)* is indicated by black with green background.

#### **4.2.5 Phylogenetic analysis of the *TaRAD51* and *TaDMC1* cDNA homoeologues**

To determine the level of conservation and evolutionary relationships between the *TaRAD51* and *TaDMC1* gene homoeologues across a diverse range of species, previously annotated database entries (Table 4.5) in addition to currently cloned *TaRAD51* and *TaDMC1* cDNA homoeologues were used to construct a neighbour-joining tree. The phylogenetic tree in Figure 4.11 shows the three main branches, with RAD51 and DMC1 members clustered in one branch, RAD51B, RAD51C, RAD51D, XRCC2 and XRCC3 into another branch and the prokaryotic RecA falling independently into a separate third branch. This indicates that RAD51 and DMC1 were paralogues and are likely to be descendants of the ancestral RecA gene. As expected the three cDNA homoeologous of *TaRAD51* and *TaDMC1* were clustered together and fall into their respective branches. The strongest similarity was found between the proteins of wheat and rice RAD51 and DMC1 cDNA sequences as would be expected due to their evolutionary relatedness.

Further analysis of RAD51 and DMC1 branches in the phylogenetic tree revealed three kingdoms, with a clear distinction between plant, animal and fungal sequences. While in the plant RAD51 and DMC1 cluster, monocotyledonous plant sequences can be distinguished from the dicotyledonous plant sequences. This is more evident in the RAD51 clade than the DMC1 because of large number of RAD51 gene entries in the database.





**Figure 4.11** Phylogenetic tree obtained from deduced amino acid sequences of the *TaRAD51* and *TaDMC1* cDNA homoeologues with a wide range of RAD51 and DMC1 gene paralogues using the neighbour-joining (NJ) method. Symbols used in the tree are: open triangle, plant RAD51; closed triangle, animal RAD51; closed grey triangle, yeast RAD51; open circle, plant DMC1; closed circle, animal DMC1; open grey circle, yeast DMC1; open diamond, plant RAD51B; closed black diamond, animal RAD51B; open inverted triangle, plant RAD51C; closed black triangle, animal RAD51C; open squares, plant RAD51D; closed black squares, animal RAD51D; XRCC2, XRCC3 and RecA are without symbols.

**Table 4.5** Annotated nucleotide sequences used for phylogenetic analysis (Coding sequences were translated and aligned in the program MEGA 4.0)

Meiotic gene	Accession number	Common name	Source
<b><i>RecA</i></b>			
<i>EcRecA</i>	V00328	<i>E. Coli RecA</i>	NCBI
<b><i>DMC1</i></b>			
<i>OsDMC1A</i>	AB079873	Rice <i>DMC1A</i>	NCBI
<i>OsDMC1B</i>	AB079874	Rice <i>DMC1B</i>	NCBI
<i>AtDMC1</i>	NM_113188	Arabidopsis <i>DMC1</i>	NCBI
<i>TaDMC1</i>	EU915561	Wheat <i>DMC1</i>	NCBI
<i>HvDMC1</i>	AF234170	Barley <i>DMC1</i>	NCBI
<i>SbDMC1</i>	Sb04g008730	Sorghum <i>DMC1</i>	Gramene
<i>SbDMC1</i>	Sb08g001020	Sorghum <i>DMC1</i>	Gramene
<i>VvDMC1</i>	GSVIVG0002006500	Grape <i>DMC1</i>	Gramene
<i>GmDMC1</i>	U66836	Soybean <i>DMC1</i>	NCBI
<i>HsDMC1</i>	NM_007068	Human <i>DMC1</i>	NCBI
<i>ScDMC1</i>	M87549	Yeast <i>DMC1</i>	
<i>MmDMC1</i>	NM_010059	Mouse <i>DMC1</i>	NCBI
<b><i>RAD51</i></b>			
<i>OsRAD51A1</i>	AB080262	Rice <i>RAD51A1</i>	NCBI

<i>OsRAD51A2</i>	AB080264	Rice <i>RAD51A2</i>	NCBI
<i>AtRAD51</i>	NM_122092	Arabidopsis <i>RAD51</i>	NCBI
<i>ZmRAD51A</i>	AF079428	Maize <i>RAD51A1</i>	NCBI
<i>ZmRAD51B</i>	AF079429	Maize <i>RAD51A2</i>	NCBI
<i>TaRAD51A1</i>	EU915557	Wheat <i>RAD51A1</i>	NCBI
<i>TaRAD51A2</i>	EU915558	Wheat <i>RAD51A2</i>	NCBI
<i>SbRAD51</i>	Sb02g037320	Sorghum <i>RAD51</i>	Gramene
<i>SbRAD51</i>	Sb05g024565	Sorghum <i>RAD51</i>	Gramene
<i>LeRAD51</i>	U22441	Tomato <i>RAD51</i>	NCBI
<i>PnRAD51</i>	AB269815	Poplar <i>RAD51</i>	NCBI
<i>HsRad51</i>	NM_002875	Human <i>RAD51</i>	NCBI
<i>MmRAD51</i>	D13803	Mouse <i>RAD51</i>	NCBI
<i>DmRad51</i>	NM_079844	Fruit fly <i>RAD51</i>	NCBI
<i>GgRad51</i>	NM_205173	Chicken <i>RAD51</i>	NCBI
<i>OcRAD51</i>	NM_001082024	Rabbit <i>RAD51</i>	NCBI
<i>ScRAD51</i>	YER095W	Yeast <i>RAD51</i>	NCBI
<b><i>RAD51B</i></b>			
<i>AtRAD51B</i>	NM_128416	Arabidopsis <i>RAD51B</i>	NCBI
<i>HsRAD51B</i>	NM_133510	Human <i>RAD51B</i>	NCBI
<i>MnRAD51B</i>	NM_009014	Mouse <i>RAD51B</i>	NCBI
<b><i>RAD51C</i></b>			
<i>AtRAD51C</i>	NM_130091	Arabidopsis <i>RAD51C</i>	NCBI
<i>TaRAD51C</i>	EU915559	Wheat <i>RAD51C</i>	NCBI
<i>HsRAD51C</i>	NM_058216	Human <i>RAD51C</i>	NCBI
<i>MnRAD51C</i>	NM_053269	Mouse <i>RAD51C</i>	NCBI
<b><i>RAD51D</i></b>			
<i>AtRAD51D</i>	NM_100649	Arabidopsis <i>RAD51D</i>	NCBI
<i>TaRAD51D</i>	EU915560	Wheat <i>RAD51D</i>	NCBI
<i>HsRAD51D</i>	AF034956	Human <i>RAD51D</i>	NCBI
<i>MnRAD51D</i>	AF034955	Mouse <i>RAD51D</i>	NCBI
<b><i>XRCC2</i></b>			
<i>AtXRCC2</i>	NM_001037064	Arabidopsis <i>XRCC2</i>	NCBI
<b><i>XRCC3</i></b>			
<i>AtXRCC3</i>	NM_125127	Arabidopsis <i>XRCC3</i>	NCBI
<i>HsXRCC3</i>	NM_005432	Human <i>XRCC3</i>	NCBI
<b><i>Lim15</i></b>			
<i>LiLIM15</i>	D21821	Lily <i>LIM15</i>	NCBI

#### **4.2.6 Comparative protein analysis of the *TaRAD51* and *TaDMC1* cDNA homoeologues**

A comparative protein analysis of TaRad51 cDNA homoeologues revealed that the three cDNA homoeologues of the TaRad51 were highly similar except for one amino acid difference between the three cDNA homoeologues in  $\alpha$ -helix 1 at the N-terminal end. Conservation was also strong with regards to the structural motifs and catalytic domains which are spaced identically across these entire homoeologous gene. The Walker A and Walker B motifs of TaRad51 cDNA homoeologues were also exactly identical (Figure 4.12).

Sequences from the three cDNA homoeologues of TaDmc1 revealed very high levels of similarity to one another for all the domains analyzed. The Walker A and Walker B motifs of TaDmc1 cDNA homoeologues were also exactly identical (Figure 4.13).

This uniformity in structural features for *TaRAD51* and *TaDMC1* cDNA homoeologous suggests that the protein structure was also well conserved and is likely to be under selective pressure.

**N-terminal domain**

		Hairpin residues		Conserved Gly residue	
		$\alpha$ -helix 1		$\alpha$ -helix 2	
					$\alpha$ -helix 3
TaRAD51A	27	PIEHLQAS	GIA	AVDVKKLK	DAGLCT VESVAY 58
TaRAD51B	27	PIEHLQAS	GIA	AVDVKKLK	DAGLCT VESVAY 58
TaRAD51D	27	PIEHLQAS	GIA	AVDVKKLK	DAGLCT VESVAY 57

		Hairpin residues			
		$\alpha$ -helix 4		$\alpha$ -helix 5	
TaRAD51A	62	KDLLQIK	GIS	EAKVDKIIEAASKL	
TaRAD51B	62	KDLLQIK	GIS	EAKVDKIIEAASKL	
TaRAD51D	61	KDLLQIK	GIS	EAKVDKIIEAASKL	

**Linker region**

		$\alpha$ -helix 6
TaRAD51A	93	ATQLHAQRL
TaRAD51B	93	ATQLHAQRL
TaRAD51D	92	ATQLHAQRL

**C-terminal domain**

		Walker A	Walker B
TaRAD51A	132	GEFRSGKT-- [84] --	LMVID226
TaRAD51B	132	GEFRSGKT-- [84] --	LMVID226
TaRAD51D	131	GEFRSGKT-- [83] --	LMVID225

**Figure 4.12** Comparative amino acid sequence analysis between the three TaRAD51 cDNA homoeologues. The amino acid residue that is different between the three cDNA homoeologues is indicated by red colour. Within the *N*-terminal domain, five  $\alpha$ -helices form two consecutive helix-hairpin-helix (HhH) motifs ( $\alpha$ -helix 1-hairpin residues-  $\alpha$ -helix 2 and  $\alpha$ -helix 4-hairpin residues-  $\alpha$ -helix 5) connected together by an  $\alpha$ -helix ( $\alpha$ -helix 3). The conserved polar non-charged glycine residue (green colour) is indicated with a red arrow. The linker region consists of a nine amino acid  $\alpha$ -helix, and the *C*-terminal domain contains the Walker A and Walker B motifs in all sequences.

**N-terminal domain**

		Hairpin residues		Conserved Gly residue			
		$\alpha$ -helix 1		$\alpha$ -helix 2		$\alpha$ -helix 3	
		↓			↓		
TaDMC1A	31	IDKLIS	GIN	SGDVKKLQD	DAGIYT	CNGLMM	60
TaDMC1B	31	IDKLIS	GIN	SGDVKKLQD	DAGIYT	CNGLMM	60
TaDMC1D	31	IDKLIS	GIN	SGDVKKLQD	DAGIYT	CNGLMM	60

		Hairpin residues			
		$\alpha$ -helix 4		$\alpha$ -helix 5	
		↓			
TaDMC1A	62	KKSLTGIR	GIR	EARVDRICEAAEKLL	
TaDMC1B	62	KKSLTGIR	GIR	EARVDRICEAAEKLL	
TaDMC1D	62	KKSLTGIR	GIR	EARVDRICEAAEKLL	

**Linker region**

		$\alpha$ -helix 6	
		↓	
TaDMC1A	94	GSDLLIR	
TaDMC1B	94	GSDLLIR	
TaDMC1D	94	GSDLLIR	

**C-terminal domain**

		Walker A	Walker B
		↓	
TaDMC1A	132	GEFRSGKT-- [83] --	LLVID228
TaDMC1B	132	GEFRSGKT-- [83] --	LLVID228
TaDMC1D	132	GEFRSGKT-- [83] --	LLVID228

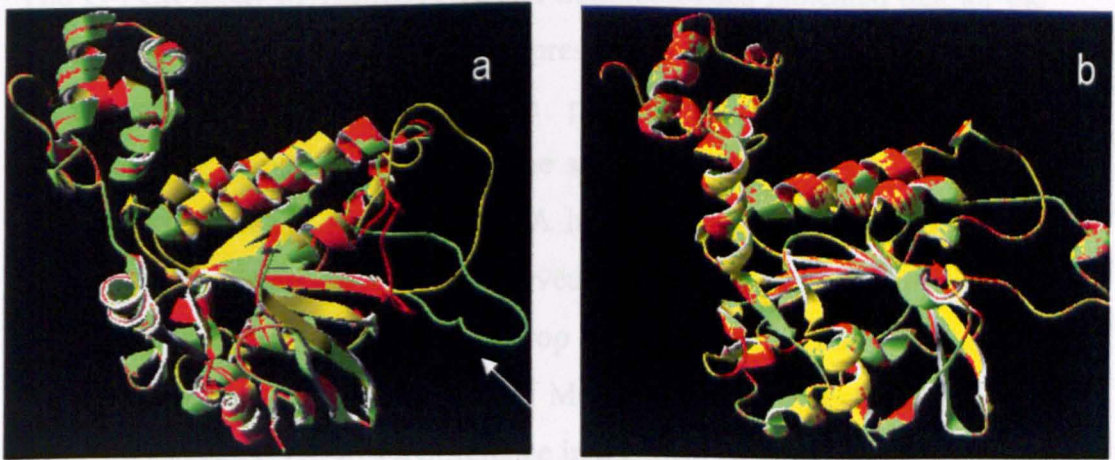
**Figure 4.13** Comparative amino acid sequence analysis between the three TaDMC1 cDNA homoeologues. Within the *N*-terminal domain, five  $\alpha$ -helices form two consecutive helix-hairpin-helix (HhH) motifs ( $\alpha$ -helix 1–hairpin residues– $\alpha$ -helix 2 and  $\alpha$ -helix 4–hairpin residues– $\alpha$ -helix 5) connected together by an  $\alpha$ -helix ( $\alpha$ -helix 3). The conserved polar non-charged glycine residue (green colour) is indicated with a red arrow. The linker region consists of a nine amino acid  $\alpha$ -helix, and the C-terminal domain contains the Walker A and Walker B motifs in all sequences.

#### **4.2.7 Structural conservation of the protein sequences of the *TaRAD51* and *TaDMC1* cDNA cDNA homoeologues**

3D molecular modeling of the TaRad51 and TaDmc1 cDNA homoeologues sequences was done to show the level of structural conservation at the protein level.

For TaRad51 cDNA homoeologues, the predicted 3D overlays superimposed onto each other revealed that all the three cDNA homoeologues were conserved and there is a high level of conservation for secondary ( $\alpha$ -helices and anti-parallel  $\beta$ -strands) and tertiary structures as well. The only noticeable structural dissimilarity observed between the three cDNA homoeologues of TaRAD51 is in peptide loops; seen between  $\alpha$ -helix 13 and  $\alpha$ -helix 14 (represented by white arrow in Figure 4.14a). The predicted 3D overlays for the TaDmc1 cDNA homoeologues superimposed onto each other revealed very high levels of conservation for secondary ( $\alpha$ -helices and anti-parallel  $\beta$ -strands) and tertiary structures as well (Figure 4.14b)





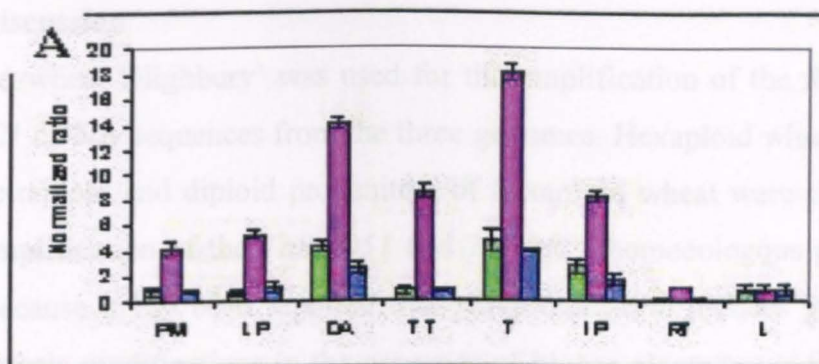
**Figure 4.14** 3D predictive modeling of TaRad51 (a) and Tadmcl (b) cDNA homoeologues superimposed on to each other. The 3D structure of TaRad51(A) is represented in red, TaRad51(B) in green and TaRad51(D) in yellow in figure 4.14a and the 3D structure of TaDmcl(A) is represented in red, TaDmcl(B) in green and TaDmcl(D) in yellow in figure 4.14b. White arrow indicated the major structural dissimilarity between the three TaRad51 cDNA homoeologues.

#### 4.2.8 Homoeologue-specific expression patterns of *TaRAD51* and *TaDMC1* cDNA homoeologues

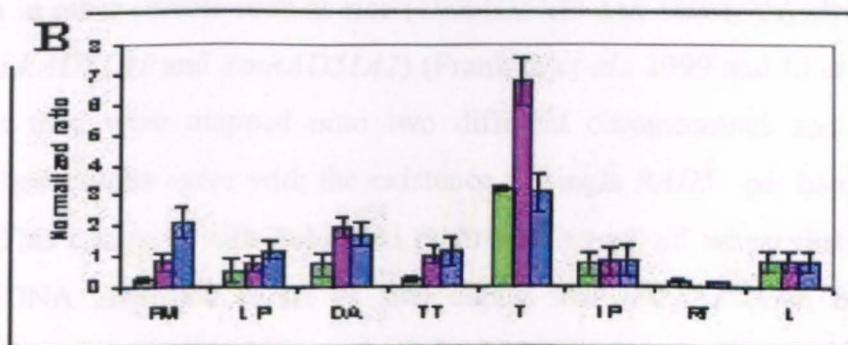
Q-PCR was undertaken to analyze the mitotic and meiotic expression patterns of *TaRAD51* and *TaDMC1* cDNA homoeologues using eight different staged tissues from wild type wheat. The tissue stages include root tip, mature leaf and meiosis specific stages: Pre-meiosis, Leptotene-Pachytene, Diplotene-AnaphaseI, TelophaseI-TelophaseII, Tetrads and Immature pollen. The primers used were the same primers that were used to isolate the different cDNA homoeologues of *TaRAD51* and *TaDMC1* (Table 4.2). The specificity of primers was again confirmed by Nulli-tetrasomic analysis and the amplification efficiency of all primers was tested before actual assay. Finally the expression of each gene was normalized against Tubulin gene expression using Tubulin primers (Table 4.2).

The Q-PCR assay of *TaRAD51* cDNA homoeologues indicated that all the three cDNA homoeologues were expressed in both mitotic and meiotic tissues and the overall expression patterns of all three *TaRAD51* homoeologues remained essentially the same (Figure 4.15A). Even though the expression of the *TaRAD51* cDNA homoeologues was seen in mitotic tissues, higher expression was observed in meiotic stages (Figure 4.15, stages PM to TT). There is a sharp drop in expression of *TaRAD51* cDNA homoeologue transcripts at the end of Meiosis I but they increase again in the Tetrad stage. There was a difference in the level of transcript of the three cDNA homoeologues, particularly during the meiotic stages which may suggests that they could have different roles in meiotic recombination. The expression of *TaRAD51(B)* was significantly higher than *TaRAD51(A)* and *TaRAD51(D)* in all meiotic stages suggesting it could be the version with the greatest role in meiosis.

The Q-PCR assay of *TaDMC1* cDNA homoeologues indicated that all the three cDNA homoeologues were expressed in both mitotic and meiotic tissues and the overall expression patterns of all *TaDMC1* cDNA homoeologues remained essentially the same (Figure 4.15B). However higher expression was found in meiotic stages (Figure 4.16, stages PM to TT) compared to mitotic tissues. There is rather a sharp spike in expression of all three *TaDMC1* cDNA homoeologues during the Tetrad stage. There was a difference in the amount of transcript of the three homoeologous genes of *TaDMC1* in the observed tissues, which may suggest that they could have different roles in meiotic recombination. Both *TaDMC1(B)* and *TaDMC1(D)* were highly expressed in all meiotic stages, except for Tetrad stage, (where expression of *TaDMC1(A)* was relatively higher compared to other stages) compared to the transcript level of *TaDMC1(A)* suggesting they could be the version with the greatest role in meiosis.



**Figure 4.15A** Expression analysis of the *TaRAD51* cDNA homoeologues by Q-PCR. The green bars represents *TaRAD51-7A*, pink bar represents *TaRAD51-7B* and blue bar represents *TaRAD51-7D*. The meiotic time course was according to Crismani *et al.*, (2006). Abbreviations used: PM, pre-meiotic interphase; LP, leptotene-pachytene, DA, diplotene-anaphase I; TT, telophase I- telophase II; T, tetrad; IP, immature pollen; RT, root tips; L, leaves.



**Figure 4.16B** Expression analysis of *TaDMC1* cDNA homoeologues by Q-PCR. The green bars represents *TaDMC1-7A*, pink bar represents *TaDMC1-7B* and blue bar represents *TaDMC1-7D*. The meiotic time course was according to Crismani *et al.*, (2006). Abbreviations used: PM, pre-meiotic interphase; LP, leptotene-pachytene, DA, diplotene-anaphase I; TT, telophase I- telophase II; T, tetrad; IP, immature pollen; RT, root tips; L, leaves.

### 4.3 Discussion

Hexaploid wheat ‘Highbury’ was used for the amplification of the *RAD51* and *DMC1* cDNA sequences from the three genomes. Hexaploid wheat and not the tetraploid and diploid progenitors of hexaploid wheat were chosen for the amplification of the *TaRAD51* and *TaDMC1* homoeologous genes. This is because it has been reported that polyploidization induces genetic and epigenetic modifications in the genomes of higher plants (reviewed by Comai, 2000). Hence the sequences cloned from the tetraploid and diploid wheat may not represent the actual sequences or the patterns of expression of those genes in hexaploid wheat may be different.

#### 4.3.1 The *RAD51* in hexaploid wheat

Nulli-tetrasomic analysis revealed that hexaploid wheat has three *TaRAD51* cDNA homoeologues located on Group 7 chromosomes with a copy on each of the three genomes. This study found that there was only one copy of *TaRAD51* per haploid genome. The duplication of the *RAD51* gene have been seen in other cereals such as rice (*OsRAD51A1* and *OsRAD51A2*) and maize (*ZmRAD51A1* and *ZmRAD51A2*) (Franklin *et al.*, 1999 and Li *et al.*, 2007) but they were mapped onto two different chromosomes and the current wheat results agree with the existence of single *RAD51* per haploid genome. This contrasts with published reports in hexaploid wheat that the *RAD51* cDNA sequence exists as two copies *TaRAD51A1* (Gen bank accession number EU915557) and *TaRAD51A2* (Gen bank accession number EU915558), both residing on the same chromosome Group 7 (Khoo *et al.*, 2008). They suggested that these two cDNA sequences might be allelic variants of the same gene from two out of the three different genomes in wheat and may not represent two highly-similar yet distinguishable genes. The current study does not agree with the results of Khoo *et al.*, (2008). We found that the reported *TaRAD51A1* is actually the (D) homoeologue and *TaRAD51A2* is the truncated (A) homoeologue of *TaRAD51*.



There is a high level of conservation of coding nucleotides, exon/intron structure and regulatory motifs among the cDNA homoeologues of *TaRAD51*. The similarity among the cDNA homoeologues of *TaRAD51* was 97% at the nucleotide level and reached 99% when comparing the proteins. Similar levels of conservation were found for the *Mre11* gene isolated from all the three progenitor genomes of wheat using *Triticum monococcum* (genome A) and *Ae. tauschii* (genome D) and the tetraploid *Triticum turgidum* (genomes A and B) (De Bustos *et al.*, 2007). This strong conservation might suggest the important role played by the *TaRAD51* cDNA homoeologues and possibly other genes involved in the process of homologous recombination and repair. That all three cDNA homoeologues of *TaRAD51* continue to be expressed in hexaploid wheat may add weight to this argument.

The Phylogenetic analysis of the *TaRAD51* homoeologous genes in this study is according to the method of Lin *et al.*, (2006). Phylogenetic analysis revealed that the cDNA homoeologues of *TaRAD51* were very similar to rice and maize *RAD51s* which suggests they are the orthologues of the rice and maize genes.

All three cDNA homoeologues of *TaRAD51* were found to code for highly conserved N-terminal, linker and C-terminal domains. Since there is a deletion of one amino acid (Glutamic acid) in the (D) homoeologue, the numbering of amino acids differed within the domains. Apart from one deletion and few substitutions, all the three cDNA homoeologues of *TaRAD51* are very identical. Miller *et al.*, (2004) reported that even a very short deletion at the N-terminal can severely disturb the folding of the Rad51 paralogues. But how far the deletion of one amino acid could have affected the overall protein structure of *TaRAD51(B)* compared to other two needs further investigation. Zhang *et al.*, (2005) found that substitution of a conserved polar non-charged glycine residue at position 103 in *ScRAD51* with a negatively-charged glutamic acid residue leads to the loss of *RAD51* DNA-binding ability. This glycine residue was found be conserved in all the three homoeologous genes of *TaRAD51* sequences analyzed. SIFT analysis

was undertaken to find out if the substitution of amino acids has any effect on the protein structure. However all the amino acid substitutions are highly tolerable indicating that there is no affect on protein structure. In addition to this amino acid, the most important motifs Walker A and B (Walker *et al.*, 1982) in the C-terminal region which give RecA both its ATP-binding and hydrolysis abilities allowing it to work as ATPase, were conserved and identical spaced for all the three homoeologous genes of TaRAD51. Analysis of *TaRAD51* cDNA homoeologues and their predicted secondary structures revealed that these proteins contain five  $\alpha$ -helices in the N-terminal domain that are found to stack upon each other to form consecutive HhH motifs, also known as (HhH)<sub>2</sub> domain. 3D predictive protein modeling with *TaRAD51* homoeologous proteins superimposed onto each other indicated that the secondary and tertiary structures of these homoeologous genes share many similar features. As just only one amino acid differences among the cDNA homoeologues of *TaRAD51* was in the important motif, no significant difference were found in the final secondary and tertiary structures of these proteins.

Q-PCR is widely used as a tool for the quantification of mRNA in basic molecular research, with applications in bio-medicine and biotechnology (Bustin, 2002). This technique has been used in many plant studies and in wheat (De Bustos *et al.*, 2007). One important feature of Q-PCR is the use of an endogenous control for the normalization of the RNA content between the samples and across tissues. For the current study  $\beta$ -Tubulin was used as the endogenous control because it was reported to show limited variation during Q-PCR experiments (Crismani *et al.*, 2007). Very few studies have involved comparing the relatively expression of homoeologous genes in hexaploid wheat (Loukoianov, 2005; Shitsukawa *et al.*, 2007a and 2007b; Bustos *et al.*, 2007). Q-PCR expression analysis results of the *TaRAD51* cDNA homoeologues in the current study were compared with microarray expression profiling of *TaRAD51* in Chapter 3. Because the Affymetrix GeneChip used in the microarray experiment is a discovery chip and does not differentiate between wheat homoeologues, most of the oligos for *TaRAD51* are found to cross-hybridize to all three cDNA homoeologues. So

the microarray expression profile of *TaRAD51* will have contributions from all the three cDNA homoeologues and the displayed expression could be an overall effect from all the three cDNA homoeologues. The current study has attempted to split the observed microarray expression patterns of *TaRAD51* gene into Q-PCR expression patterns of individual cDNA homoeologues. Q-PCR expression analysis of the *TaRAD51* cDNA homoeologues indicated that all the three cDNA homoeologues are expressed in both vegetative and reproductive stages but they are up-regulated during the reproductive phase. Similar expression patterns were observed in microarray expression profiling of *TaRAD51* transcripts (Chapter 4) and published in reports from Arabidopsis, rice and wheat (Doutriaux *et al.*, 1998; Ding *et al.*, 2001; Osakabe *et al.*, 2002; Khoo *et al.*, 2008). However, the one contrasting difference observed in the current Q-PCR result with that of microarray expression results by Khoo *et al.*, (2008) in wheat is the elevated expression levels of one of the three *TaRAD51* cDNA homoeologues (*TaRAD51(B)*) during the Tetrad stage of meiosis. This needs to be further investigated. Because the combined Q-PCR results of *TaRAD51* cDNA homoeologues from the current study correlate with that of the microarray expression pattern in Chapter 3, it can be concluded that the microarray expression profile of *TaRAD51* will have contributions from all the cDNA homoeologues.

#### 4.3.2 The *DMC1* in hexaploid wheat

Hexaploid wheat has three *TaDMC1* cDNA homoeologues located on Group 5 chromosomes with a copy on each of the three genomes as revealed by Nulli-tetrasomic analysis. This study found that there was only one copy of *TaDMC1* per haploid genome. The duplication of the *DMC1* genes has been seen in other cereals such as rice (*OsDMC1A* and *OsDMC1B*) (Ding *et al.*, 2001). But no duplication of the *DMC1* genes was observed in hexaploid wheat, per genome. The current study also found that the *TaDMC1* cDNA sequence reported by Khoo *et al.*, (2008) is actually the *D* homoeologue of *TaDMC1* gene. There is a high level of conservation of coding nucleotides, exon/intron structure and regulatory motifs among the



homoeologous genes of *TaDMC1*. The similarity among the cDNA homoeologues of *TaRAD51* was 97% at the nucleotide level and reached 99%, when comparing the proteins. This strong conservation might suggest an important role played by *TaDMC1* cDNA homoeologues. That all three cDNA homoeologues of *TaDMC1* continue to be expressed in the hexaploid wheat may add weight to this argument in a similar way to the expression of all *TaRAD51* cDNA homoeologues in this study. Phylogenetic analysis revealed that the cDNA homoeologues of *TaDMC1* were very similar to rice and barley *DMC1* which suggests they are the orthologues of the rice and barley genes.

Apart from a few amino acid substitutions, all three *TaDMC1* cDNA homoeologues were found to code for highly conserved N-terminal, linker and C-terminal domains. SIFT analysis was undertaken to find out if the substitution of amino acids has any effect on the protein structure. However all the amino acid substitutions are highly tolerable indicating that there is no affect on protein structure. The glycine residue was found be conserved in all the three homoeologous genes of *TaDMC1* sequences analyzed. In addition to this amino acid, the most important motifs Walker A and B (Walker *et al.*, 1982) in the C-terminal region are conserved and identically spaced for all the three cDNA homoeologues of *TaDMC1*. Analysis of the *TaDMC1* cDNA homoeologues and their predicted secondary structures revealed that these proteins contain five  $\alpha$ -helices in the N-terminal domain that are found to stack upon each other to form consecutive HhH motifs, also known as (HhH)<sub>2</sub> domain. 3D predictive protein modeling with *TaDMC1* cDNA homoeologues superimposed onto each other indicated that the secondary and tertiary structures of these homoeologous proteins share many similar features. No difference was found in the final secondary and tertiary structures of these proteins, with the only major predicted changes being in linker regions, as for the *TaRAD51* homoeologues.

Q-PCR expression analysis of the *TaDMC1* cDNA homoeologues indicated that all the three are expressed in both vegetative and reproductive stages and they are up-regulated during meiosis. Contrasting differences were observed in the current Q-PCR study with that of the microarray expression and results by Khoo *et al.*, (2008) during the Tetrad stage where elevated expression patterns of all the three *TaDMC1* cDNA homoeologues was observed. This needs to be further investigated. Because the Q-PCR results of only one of the three *TaDMC1* cDNA homoeologues from the current study correlate with that of the microarray expression pattern, it can be concurred that the microarray expression profile of *TaRAD51* gene will have contributions from only one out of the three cDNA homoeologues. This can be explained on the basis that only the two oligos corresponding to the *TaDMC1(D)* (Chapter 4) will carry maximum signal intensities compared to the rest of 8 oligos.

In summary, all the three cDNA homoeologues of *TaRAD51* and *TaDMC1* have been cloned from hexaploid wheat based on sequence differences identified. Nucleotide alignments, amino acid alignments and intron/exon structure indicated that all the three cDNA homoeologues of *TaRAD51* and *TaDMC1* are identical with very few substitutions of amino acids and one insertion/deletion of an amino acid (for the *TaRAD51(D)* homoeologue). Phylogenetic analysis of the three cDNA homoeologues of *TaRAD51* and of *TaDMC1* with the eukaryotic *RAD51* and *DMC1* showed that they are the true orthologues of other cereals strand-exchange proteins, such as rice, maize and barley. 3D predictive protein modeling of *TaRAD51* and *TaDMC1* homoeologous proteins superimposed onto the corresponding homoeologues indicated that the secondary and tertiary structures of these homoeologous genes share many similar features despite minor changes. Finally Q-PCR expression profiling experiments revealed that there are differences in the expression levels of the three cDNA homoeologues of *TaRAD51* and *TaDMC1*, possibly reflecting a difference in their roles in homologous recombination. But overall relative patterns of expression remained same for all the three cDNA homoeologues of *TaRAD51* and

*TaDMC1* genes, respectively. Whether these differences have any functional significance will be pursued in the next chapter.

## **CHAPTER 5: FUNCTIONAL CHARACTERISATION OF THE *RAD51* AND *DMC1* GENES IN WHEAT (*Triticum aestivum* L.)**

### **5.1 Introduction**

Forward and/or Reverse Genetics are the two main approaches that are used to determine the function of a particular gene/genes through screening the phenotype or genotype of individual mutants. In the Forward Genetics approach, one starts with a particular identified phenotype or biological process (such as flower colour or ear fertility or plant height) and the sequence of the particular gene responsible is ultimately deduced, often through the genetic and phenotypic analysis of large numbers of mutagenized individuals. This approach is particularly useful when no sequence information is available for the gene of interest or for related genes in other species. Forward Genetic approaches are not practical for genome wide analysis in many crops because they are labour intensive and time consuming to screen large numbers of mutagenized plants to identify each gene coding for a particular phenotype (Alonso and Ecker, 2006). It is also less applicable for those gene/genes with no clear phenotype or with potentially quantitative phenotypes.

In the Reverse Genetics approach, one starts with the known gene sequence and mutant populations are screened to identify individuals with sequence alterations for that particular gene of interest or the associated non-coding regions, such as promoters. By contrast to the Forward Genetics approach, this can approach is less time consuming and the widespread availability of sequence data in a number of model plant species makes this approach attractive for researchers, to rapidly design Reverse Genetics strategies to functionally characterize the effects of modification of a particular gene/genes. In both Forward and Reverse Genetic approaches, the function of a particular gene/genes are typically assigned to a specific biological

process by analyzing the phenotypic consequences of altering the activity of that particular gene/genes.

A long-recognized large-scale approach for achieving random gene inactivation is to induce mutations in a population of plants using chemical or physical mutagen agents. This approach is of particular feasible because it is amenable to most plant species regardless of their transformability. When coupled with sensitive methods for the detection of 'aberrant' DNA fragments in complex PCR-derived mixtures, this previously random approach becomes amenable to the identification of mutations in targeted genes by reverse genetics.

Chemical mutagenesis is a well recognized large-scale method for achieving random DNA sequence modification in a population of plants. This approach is of particular interest because it is applicable to most plant species regardless of their transformability. This approach can easily be coupled with sensitive methods for the detection of 'aberrant' DNA fragments in complex PCR-derived mixtures for the identification of mutations in targeted genes by Reverse Genetics. EMS (Ethyl Methane Sulphonate) mutagenesis methods can cause gene disruption/modification by base substitution and this approach is relatively independent of genome size, with high levels of mutation possible. This reduces the overall population size required (McCallum *et al.*, 2000a, b). This method has been used successfully to develop two large-scale mutant populations in barley (*Hordeum vulgare* cv. Optic) to promote both Forward and Reverse Genetics (Caldwell *et al.*, 2004).

In plants, fast neutrons have been shown to be a very effective mutagen (Koornneef *et al.*, 1982). Around 2500 lines treated with fast neutron radiation at a dose of 60 Gy are required to inactivate every gene once on average in Arabidopsis, as an example (Koornneef *et al.*, 1982). As the Arabidopsis genome contains about 25,000 genes (Arabidopsis Genome Initiative, 2000), it is estimated that on average 10 genes are randomly disrupted in each line derived from the above mutagenesis treatment.

Molecular characterization of the *Arabidopsis gal-3* (Sun *et al.*, 1992) and tomato *prf-3* (Salmeron *et al.*, 1996) genes further demonstrated that fast neutron bombardment induces deletion mutations.

Insertional mutagenesis methods use DNA elements that are able to insert generally at random into chromosomes, such as transposons (Sundaresan *et al.*, 1995) or the T-DNA of *Agrobacterium tumefaciens* (Azpiroz-Lechan and Feldmann, 1997) to create loss of function mutations in plants. It is one of the most powerful methods and has been widely used to analyze gene function in mice, yeast, *E. coli*, *Arabidopsis* and rice (Koller *et al.*, 1989; Kempin *et al.*, 1997).

Hexaploid wheat (*Triticum aestivum* L.) provides an experimental challenge to researchers because of its polyploid nature. Both Forward and Reverse Genetics approaches have been tried in wheat. In general Forward Genetics approaches are more difficult in wheat than in diploid species because of the potentially redundant nature of the genes in the hexaploid making it more difficult to produce loss of function mutants that produce phenotypes. EMS induced Reverse Genetic approaches such as TILLING (Targeting induced local lesions in genomes) were shown to be suitable for wheat. Slade *et al.*, (2005) have demonstrated the use of TILLING in two commercially important crops with complex genomes: bread and durum wheats. By creating and screening two genetically diverse populations, they identified 246 modified alleles of the *waxy* gene in elite TILLING cultivars. An RNAi based Reverse Genetics strategy was successfully used by Fu *et al.*, (2007) to suppress the action of three homoeologous genes, even in polyploid wheat. In wheat the use of ionizing radiation were shown to have significant effect as it can delete large portions of the chromosome and often complete gene homoeologues, which can aid in both Forward and Reverse Genetics approaches (Roberts *et al.*, 1999).

Successful Reverse Genetics approaches for characterizing homoeologous genes in wheat depend on the development of Genome-Specific markers. Genome-Specific PCR primers have been designed for various purposes,

such as amplifying specific genomic sequences from hexaploid wheat for direct comparison between wheat and its diploid progenitors (Talbert *et al.*, 1998) and for monitoring segregating alleles in breeding populations (Blake *et al.*, 2004). The development of Genome-Specific PCR primers is more realistic if they are based on intragenic sequences rather than exonic sequences as introns and untranslated regions are generally more polymorphic than exons (Haga *et al.*, 2002). Genomic DNA amplified using exon-anchored primers which amplify across one or more introns allow the cloning and sequencing of introns and the development of allele-specific markers (e.g., Van Campenhout *et al.*, 2003). The location of introns in wheat genes can often be inferred by the alignment of wheat cDNA and EST sequences with the rice genomic sequence since the intron positions are usually highly conserved between wheat genes and those of its Triticeae relatives and rice genes (Dubcovsky *et al.*, 2001; Ramakrishna *et al.*, 2002).

Even though there are recent reports of the molecular characterization of *TaRAD51* and *TaDMC1* genes (Khoo *et al.*, 2009), there is not a single report on the *in planta* functional characterization of these genes and gene homoeologues in hexaploid wheat. In this chapter, both Forward and Reverse Genetics approaches were used to functionally characterize the *TaRAD51* and *TaDMC1* genes both of which have three copies (homoeologues) that are expressed in wheat, using two kinds of mutagenized population. The identified mutant lines were used to provisionally study the individual effects of the deletion of the homoeologous gene on observed phenotype, genotype and cytology, in comparison to the wild-type wheat cultivar.



## 5.2 Results

### 5.2.1 Design of exon-anchored primers for the *TaRAD51* and *TaDMC1* genes

Because the genomic sequence of the *TaRAD51* and *TaDMC1* homoeologous genes had not been reported in the database at the time of the design of Genome-Specific Primers, *OsRAD51A1* (Gen Bank accession number AB080261) and *OsDMC1A* (Gen Bank accession number AB079873) cDNA sequences were used to predict the exon/intron boundaries of the *TaRAD51* and *TaDMC1*. Based on this prediction both *TaRAD51* and *TaDMC1* cDNA sequences were expected to contain nine exons. Primer sets were designed anchored in exons and spanning one or more introns, based on this alignment. Primers were designed for *TaRAD51* and also for *TaDMC1* which were expected to amplify fragments of all three homoeologues at once (Table 5.1).

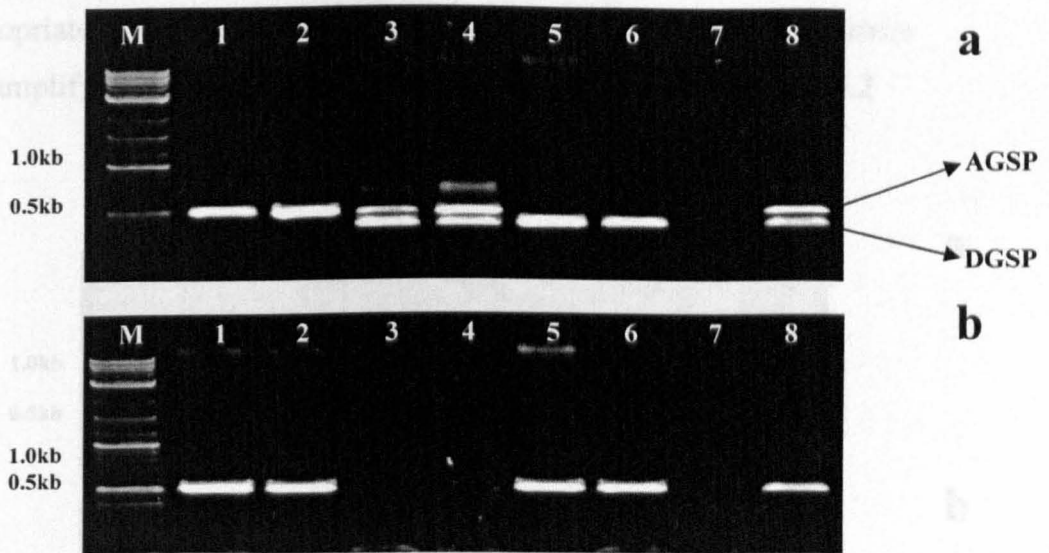
**Table 5.1** Exon-anchored oligonucleotide primer pairs used to amplify the intergenic regions of the *TaRAD51* and *TaDMC1* genes.

Gene/Location	Sequence (5'-3')	Tm (°C)	Product size (Kb)
<i>TaRAD51</i>			
Exon 8	CCATGATGGTGGAGACAAGG	65	2.5
3'UTR	TGGTGGTCCAATATCACATAGGAG		
<i>TaDMC1</i>			
5'UTR	ATGATCCACATTCCACCCGC	65	2.5
Exon 6	ATGAGCCAACTGGGTCTTCC		

## 5.2.2 Genome-Specific Primer design for the *TaRAD51* and *TaDMC1* genes

### 5.2.2.1 *TaRAD51* Genome-Specific Primers

Multiple alignments of sequences obtained from the *TaRAD51* cDNA exon-anchored primers identified three distinct groups of sequences which differed mostly in the introns. Putative primers which amplify a particular homoeologue were designed and coupled with an original exon anchored primer in a PCR reaction containing gDNA of Chinese Spring as template. As the gene had already been mapped to group 7 chromosomes on the wheat genome (Chapter 4), the same combination of primers was used to amplify the 7 group of Nulli-tetrasomic stocks in a PCR-based Nulli-tetrasomic analysis to test for the genome specificity of the primer sets. Screening with several combinations of primer sets identified one primer combination which amplified two products in Chinese Spring and in NT7B7A and NT7B7D but one product in NT7A7B, NT7A7D, NT7D7A and NT7D7B. The two product sizes in NT7B7A and NT7B7D can be assigned to chromosomes 7D and 7A based on their absence in respective NT lines (Figure 5.1). To further design primers that amplify a particular homoeologue either 7A or 7D, these PCR amplified products were cloned, sequenced and aligned to reveal a 26-bp insertion/deletion (indel) event that distinguishes the 7D and 7A sequences. Screening with another primer set amplified a product from Chinese Spring, NT7A7B, NT7A7D, NT7D7A and NT7D7B but not from NT7B7A and NT7B7D so that this primer set is B genome-specific. The list of Genome-Specific Primers used to amplify a particular *TaRAD51* homoeologue is shown in Table 5.2

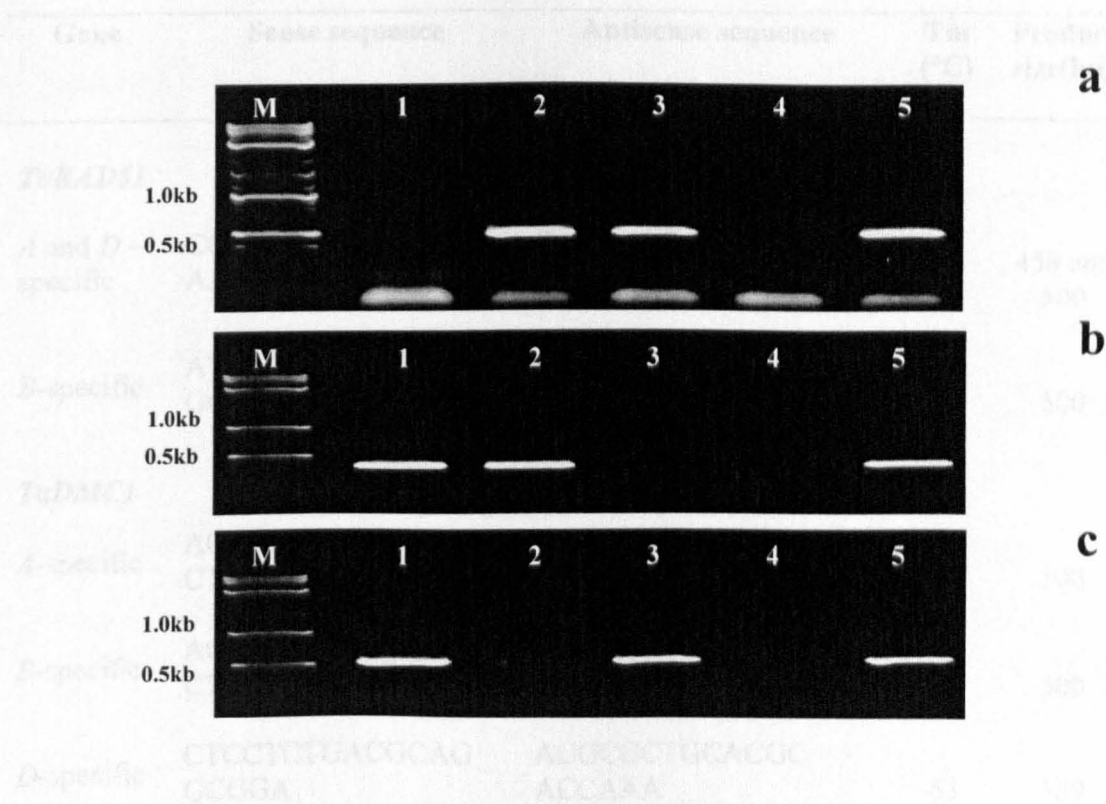


**Figure 5.1** PCR assays to test genome-specific *TaRAD51* primers. *TaRAD51* (A) and (D) genome-specific primer PCR amplification (a); The two black arrows on the right indicate *TaRAD51* (A) and (D) genome specific bands; *TaRAD51*(B) Genome-Specific Primer PCR amplification (b); 1-NT7A7B, 2-NT7A7D, 3-NT7D7A, 4-NT7DAB, 5-NT7B7A, 6-NT7B7D, 7-Water control and 8-Paragon control; M, 2-log ladder.

### 5.3.2.2 *TaDMC1* Genome-Specific Primers

Multiple alignments of sequences obtained from the *TaDMC1* cDNA exon-anchored primers identified three distinct groups of sequences with variation mostly in the introns. Putative primers which amplify a particular homoeologue were designed and coupled with an original exon-anchored primer in a PCR reaction containing gDNA of Chinese Spring as template. As the gene had already been mapped on group 5 chromosomes of the wheat genome (Chapter 4), the same combination of primers was used to amplify the 5 group of Nulli-tetrasomic stocks in a PCR-based Nulli-tetrasomic analysis to test for the genome specificity of the primer sets. Screening with several combinations of primer sets identified one primer combination for each of the homoeologues of *DMC1* gene. Each of which could be assigned to chromosomes 5A, 5B and 5D based on their absence in

the appropriate NT lines (Figure 5.2). The list of Genome-Specific Primers used to amplify a particular *TaDMC1* homoeologue is shown in Table 5.2



**Figure 5.2** PCR assays to test for genome-specific *TaDMC1* primers. *TaDMC1(A)* genome-specific primer PCR amplification (a), *TaDMC1(B)* genome-specific primer PCR amplification (b), *TaDMC1(D)* genome-specific primer PCR amplification (c); 1-NT5A5B, 2-NT5B5A, 3-NT5D5A, 4-Water control and 5-Paragon control; M, 2-log ladder

**Table 5.2** *TaRAD51* and *TaDMC1* Genome-Specific Primers used for genotyping purposes.

Gene	Sense sequence	Antisense sequence	T <sub>m</sub> (°C)	Product size(bp)
<i>TaRAD51</i>				
<i>A</i> and <i>D</i> – specific	CGGAAGGATTGGTA AAAAAT	CACTCAGAAATGACG AAAAAGG	55	450 and 500
<i>B</i> -specific	ATGGTGGAGACAA GGTGAG	GTACACTAGCATTAC GGTACAG	63	500
<i>TaDMC1</i>				
<i>A</i> -specific	AGCCTCCGCCCCA CTTCCTTC	ACAAACGCAACACGA GCACACG	63	500
<i>B</i> -specific	AGCCTTGCCCCA CTTCCTC	ACGCGCTGCACGCA CCAAA	63	500
<i>D</i> -specific	CTCCTCTGACGCAG GCGGA	ACGCGCTGCACGC ACCAAA	63	389

### 5.2.3 Forward and Reverse Genetic screening of the Highbury deletion lines

#### 5.2.3.1 Estimation of mutation frequency in the Highbury deletion lines

Because the Highbury population was at a late (M<sub>7</sub>) generation, the lines were initially tested with a set of ten microsatellite primer pairs (SSR markers) sampling all the three wheat genomes (Table 5.3) to determine the frequencies of deletions left in the population. But because two of the SSR markers (CFA2141 and CFD86) are on the same chromosome amplifying two homoeologues simultaneously (5A/5D, 5B/5D respectively) and if at all there are deletion for one homoeologue, analysis will be difficult because the other homoeologues may compensate for the lost band (if they amplify the same size as the other). So for estimating the frequency of deletions and for analysis purposes, only eight markers are considered. PCR amplification of the first 200 M<sub>7</sub> plants tested over several regions of the genome with eight microsatellite primers identified no clear deletions (through consistent

loss of the microsatellite band amplification) suggesting that limited numbers of gene-specific deletions would be expected in the full set of 1060 M<sub>7</sub> lines.

**Table 5.3** Genome-specific SSR markers used for estimating the frequency of deletions in the Highbury deletion population. The sequence of the SSR primers was obtained from the Graingene database ([www.wheat.pw.usda.gov](http://www.wheat.pw.usda.gov)).

SSR Marker	Sense Primer	Antisense Primer	T <sub>m</sub> (°C)	Chromosome location
CFA2141	GAATGGAAGGCGGACATAGA	GCCTCCACAACAGCCATAAT	60	5A/ 5D
CFA2219	TCTGCCGAGTCACTTCATTG	GACAAGGCCAGTCCAAAAGA	60	1A
WMS610	CTGCCTTCTCCATGGTTTGT	AATGGCCAAAGGTTATGAAGG	60	4A
WMS4	GCTGATGCATATAATGCTGT	CACTGTCTGTATCACTCTGCT	55	4A
WMS155	CAATCATTTCCCCCTCCC	AATCATTGGAAATCCATATGCC	60	3A
WMS261	CTCCCTGTACGCCTAAGGC	CTCGCGCTACTAGCCATTG	55	2D
WMS304	AGGAAACAGAAATATCGCGG	AGGACTGTGGGAATGAATG	55	5A
WMC416	AGCCCTTTCTACCGTGTTCCTT	TATGGTCGATGGACTGTCCCTA	61	6D
CFD86	TTAATGAGCGTCAGTACTCCC	GCAACCATGTTTAAGCCGAT	60	5D/ 5B
BARC37	AGCGCTCCCCGACTCAGATCC TT	GCGCCATGTTTCTTTTATTACTC ACTTT	55	6A

### 5.2.3.2 Forward genetic screening of Highbury deletion lines

For Forward Genetics screening, a total of 21,200 Highbury seeds from 1060M<sub>7</sub> lines with 20 plants per row (each row representing one M<sub>7</sub> line) were sown in the field. Ten typical ears for each M<sub>7</sub> line were harvested and fertility assessment was done in terms of number of grains/spikelet. Out of total of 1060 M<sub>7</sub> lines, only 10 M<sub>7</sub> lines showed phenotypic differences in terms of partial sterility with the arbitrarily chosen level indicating a potentially interesting line of grains/spikelet is  $\leq 2$  when a single ear was

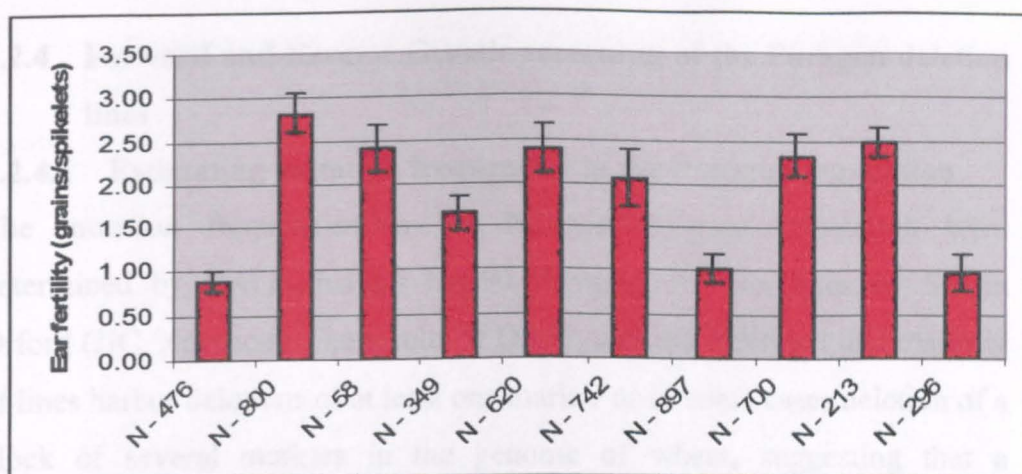
analysed. To investigate this initial result further, the remaining 9 ears collected for the lines were analysed. Significant effects on fertility across all the 10 ears were confirmed using Duncan statistical analysis for just four  $M_7$  lines (N-476, N-349, N-897 and N-296) where the average number of grains/spikelet from all 10 ears was  $\leq 2$  (Table 5.4 and Figure 5.3). All ten  $M_7$  lines were selected for further characterization even though six of the  $M_7$  lines (N-742, N-800, N-58, N-620, N-700 and N-213) had an average number of grains/spikelet  $> 2$ . These were drawn from the initial 10  $M_7$  lines and were included in the analysis and treated as controls (Table 5.4 and Figure 5.3).

**Table 5.4** Ear fertility (number of grains/spikelet-average from 10 ears) values in Highbury wheat deletion lines by a Forward Genetics approach.

Highbury deletion line number	Ear fertility (grains/spikelet)
N - 476	0.8944 $\pm$ 0.10 <sup>a</sup>
N - 800	2.8289 $\pm$ 0.23 <sup>f</sup>
N - 58	2.4133 $\pm$ 0.26 <sup>ef</sup>
N - 349	1.6622 $\pm$ 0.19 <sup>bc</sup>
N - 620	2.3978 $\pm$ 0.28 <sup>ef</sup>
N - 742	2.0644 $\pm$ 0.32 <sup>cd</sup>
N - 897	1.0344 $\pm$ 0.17 <sup>ab</sup>
N - 700	2.2956 $\pm$ 0.25 <sup>cde</sup>
N - 213	2.5000 $\pm$ 0.16 <sup>ef</sup>
N - 296	0.9833 $\pm$ 0.21 <sup>ab</sup>

Values are means of 10 replicates per treatment  $\pm$  s.e. Values for the same parameter labeled with the same letter are not significantly different ( $P \leq 0.05$ )





**Figure 5.3** Bar graph showing a comparison of 10 selected Highbury M<sub>7</sub> lines for partial ear fertility (grains/spikelet; average from 10 ears).

#### 5.3.3.2.1 Genetic analysis of the identified Highbury deletion lines identified by Forward Genetic screening

The 10 identified Highbury deletion lines were screened with the *TaRAD51* and *TaDMC1* Genome-Specific Primers (Table 5.2) to check if the observed partial sterility is due to deletions in any of the *TaRAD51* and *TaDMC1* homoeologues. PCR amplification of the ten Highbury deletion lines did not identify any line with deletions for the *TaRAD51* homoeologous genes, as might be expected, indicating that the partial sterility observed did not result from a deletion of any of the *TaRAD51* homoeologues. The 10 identified Highbury mutants with reduced fertility were also subjected to screening with *TaDMC1* Genome-Specific Primers to identify any *TaDMC1* genome specific deletions. PCR amplification of the ten Highbury deletion lines did not identify any lines with deletions for the *TaDMC1* homoeologous genes. The identified lines have been kept, but the low apparent level of deletions observed has led to the work in this population being discontinued for the moment.



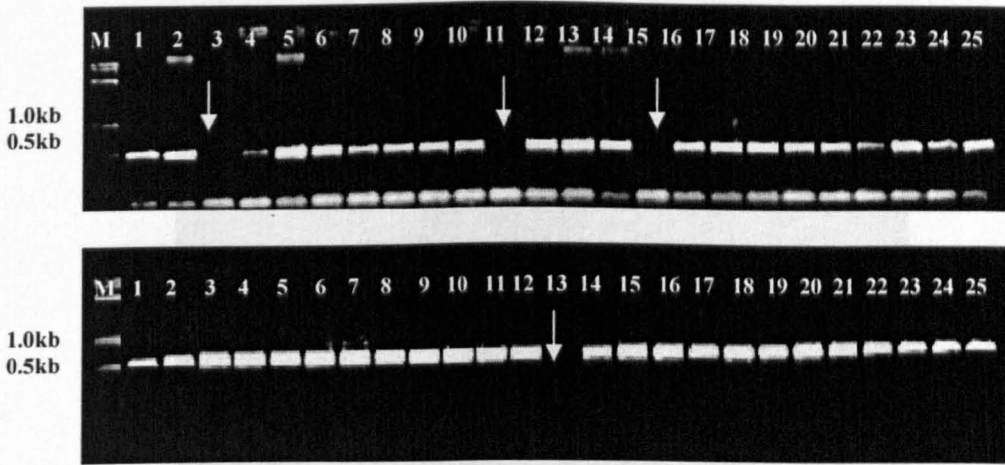
## **5.2.4 Forward and Reverse Genetic screening of the Paragon deletion lines**

### **5.2.4.1 Estimating mutation frequencies in the Paragon population**

The mutation frequencies in the Paragon deletion population were determined by DArT analysis for 90 Paragon deletion lines by Simon Orford (JIC, Norwich). The results of DArT analysis indicated that majority of lines harbor deletions of at least one marker or in some cases deletion of a block of several markers in the genome of wheat, suggesting that a reasonable number of gene-specific deletions would be expected in the full set of 450 M<sub>2</sub> lines (personal communication-Simon Griffiths, JIC, Norwich).

### **5.2.4.2 Reverse Genetics screening of the Paragon deletion population with the *TaRAD51* and *TaDMC1* Genome-Specific Primers**

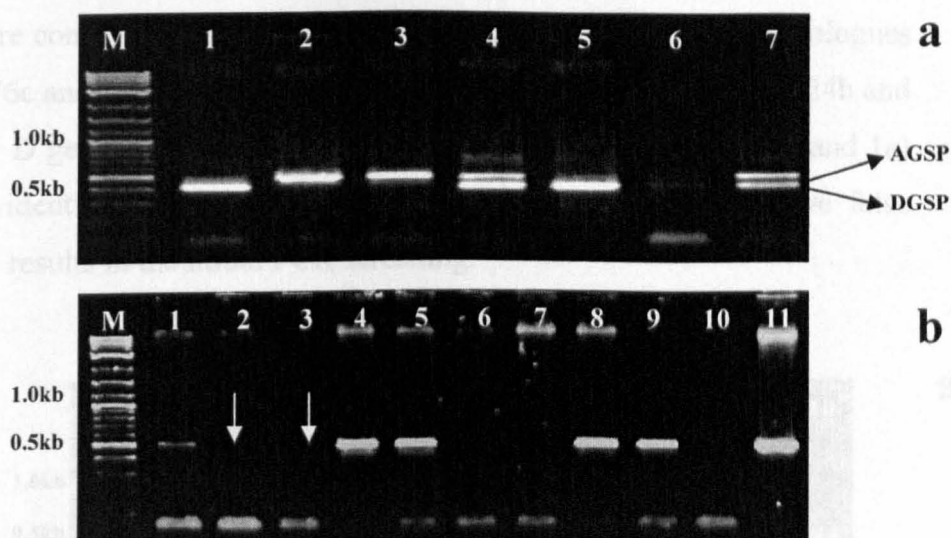
For the current experiment, ~450 M<sub>2</sub> lines of the Paragon deletion population together with unmutated Paragon were grown under glasshouse conditions at JIC, Norwich. Genomic DNA from 450 M<sub>2</sub> lines was extracted from young leaf material by me and PCR amplified with the *TaRAD51* and *TaDMC1* Genome-Specific Primers (Table 5.2). Deletions within/of any of the gene homoeologues were expected to produce either as total absence of band or a shorter band compared to the full-length PCR product in unmutated Paragon. In the current study, where a change was identified in any of these six homoeologues the total absence of band was observed, indicating that the deletion size is likely to be bigger than the GSP product in all cases. An example of the PCR assay using the *TaRAD51(B)* Genome-Specific Primers and the *TaDMC1(A)* Genome-Specific Primers with the Paragon deletion lines is shown in Figure 5.4.



**Figure 5.4** An example of mutation screening of Paragon deletion lines with the Genome-Specific Primers. PCR screening with the *TaRAD51(B)* Genome-Specific Primers (a), *TaDMC1(A)* Genome-Specific Primers (b); Arrows indicate Paragon lines with deletions identified with respective primers; M, 2-log ladder.

#### 5.2.4.2.1 Screening for deletions with *TaRAD51* Genome-Specific Primers

Preliminary PCR screening of ~450 M<sub>2</sub> lines of the paragon deletion population with the *TaRAD51* Genome-Specific Primers identified one M<sub>2</sub> line lacking the *TaRAD51(A)* gene (line 319c), three lines deleted for the *TaRAD51(B)* gene (lines 16b, 81a and 287a) and one line deleted for the *TaRAD51(D)* genome (line 221). PCR reamplification was done using gDNA from Paragon and the appropriate Nulli-tetrasomic lines (Group 7 for *RAD51*) to verify genome specificity and the presence of the deletion. Finally four M<sub>2</sub> lines were confirmed to have deletions for *TaRAD51* gene homoeologues (319c for *A* genome, line 81a and 287a for *B* genome and 221 for *D* genome) with the line 16b giving a false negative result in the initial PCR screening with *TaRAD51(B)* Genome-Specific primers (Figure 5.5).

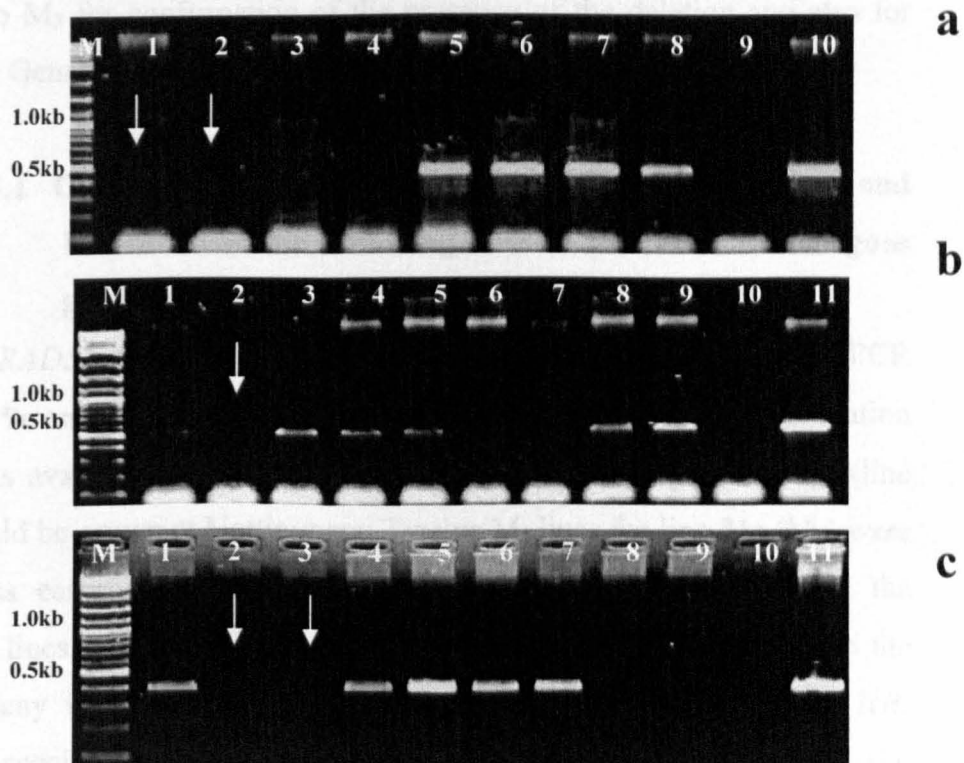


**Figure 5.11** Identification of Paragon mutant lines with *TaRAD51* Genome-Specific Primers and confirmation of the presence of a deletion using the appropriate Nulli-tetrasomic stock. *TaRAD51(A)* and (*D*) genome-specific PCR (a); 1-line 319c, 2-line 221, 3-NT7A7B, 4-NT7B7A, 5-NT7D7A, 6-Water control and 7-Paragon control; Black arrows indicates *A* Genome-Specific deletion and *D* Genome-Specific deletion; *TaRAD51(B)* genome specific PCR (b); 1-line 16b, 2-line 81a, 3-line 287c, 4-NT7A7B, 5-NT7A7D, 6-NT7B7A, 7-NT7B7D, 8-NT7D7A, 9-NT7D7B, 10-Water control and 11-Paragon control; White arrows indicate Paragon lines with *TaRAD51(B)* specific deletions only; M, 2-log ladder.

#### 5.2.4.2.2 Screening for deletions with *TaDMC1* Genome-Specific Primers

Preliminary PCR screening of ~450 M<sub>2</sub> plants of the Paragon population with *TaDMC1* Genome-Specific Primers identified two deletion lines lacking the *TaDMC1(A)* gene (lines 176c and 24b), three deletion lines for the *TaDMC1(B)* gene (lines 21b, 287a and 368c) and three deletion lines for the *TaDMC1(D)* genome (lines 1a, 24b and 260a). PCR reamplification was done using DNA from Paragon and the appropriate Nulli-tetrasomic lines (Group 5 for *DMC1* genes) to verify genome specificity. Finally five M<sub>2</sub>

lines were confirmed to have deletions for the *TaDMC1* gene homoeologues (lines 176c and 24b for *A* genome, line 287a for *B* genome and lines 24b and 260a for *D* genome) (Figure 5.11). The rest of the lines (21b, 368c and 1a) initially identified were found to be non-deletion lines which gave false negative results in the initial PCR screening.



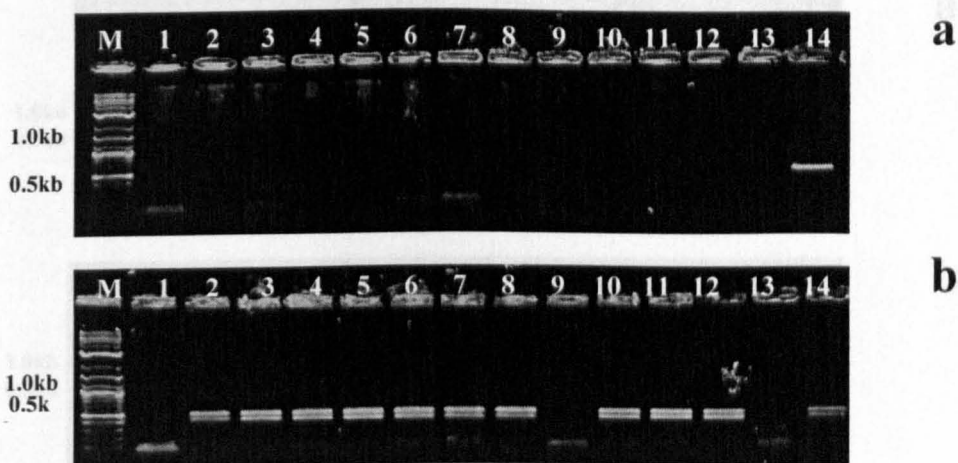
**Figure 5.6** Identification of Paragon mutant lines with *TaDMC1* Genome-Specific Primers and confirmation using appropriate Nullitetrasonic stock. *TaDMC1(A)* Genome-Specific PCR (a); 1-line 176c, 2-line 24b, 3-NT5A5B, 4-NT5A5D, 5-NT5B5A, 6-NT5B5D, 7-NT5D5A, 8-NT5D5B, 9-Water Control and 10-Paragon Control; *TaDMC1(B)* Genome-Specific PCR (b); 1- line 21b, 2-line 287a, 3-line 368c, 4-NT5A5B, 5-NT5A5D, 6-NT5B5A, 7-NT5B5D, 8-NT5D5A, 9-NT5D5B, 10-Water Control and 11-Paragon Control; *TaDMC1(D)* Genome-Specific PCR (c); 1-1a, 2-24b, 3-260a, 4-NT5A5B, 5-NT5A5D, 6-NT5B5A, 7-NT5B5D, 8-NT5D5A, 9-NT5D5B, 10-Water Control and 11-Paragon Control; White arrows indicate Paragon lines with deletions only; M, 2-log ladder.

#### **5.2.4.2.3 Forward genetic screening of M<sub>3</sub> progeny of the confirmed *TaRAD51* and *TaDMC1* homoeologous gene mutants**

Out of six confirmed M<sub>2</sub> deletion lines for both *TaRAD51* and *TaDMC1* genes, only three lines: *TaRAD51(B)* (line 81a), *TaDMC1(A)* (line 176c) and *TaDMC1(D)* (line 260a) survived probably because of heavy mutation load within this population. The seed obtained from these M<sub>2</sub> lines has been grown to M<sub>3</sub> for confirmation of the presence of the deletion and also for Forward Genetics screening.

##### **5.2.4.2.3.1 Confirmation of the transmission of the deletion and Forward Genetic screening of the *TaRAD51* homoeologous gene mutants**

Both *TaRAD51(A)* and *(D)* homoeologous gene mutants identified by PCR (lines 319c and 221) did not survive and no seed from previous generation (M<sub>1</sub>) was available. So only the *TaRAD51(B)* homoeologous mutant (line 81a) could be grown at Nottingham. Twelve M<sub>3</sub> lines for line 81a (M<sub>2</sub>) were grown as ear rows under glasshouse conditions. Before subjecting the deletion lines to Forward Genetics, the transmission of the deletion in the M<sub>3</sub> progeny was confirmed by PCR amplification with the *TaRAD51(B)* genome-specific primer and as a control with *TaRAD51(A)* and *(D)* Genome-Specific Primers (Table 5.2). As expected, the PCR amplifications indicated that the deletion was transmitted to all M<sub>3</sub> progeny (Figure 5.7). Visible phenotypes, if there were any, were scored regularly during the growing season including key developmental time points for all the M<sub>3</sub> progeny in the glass house.

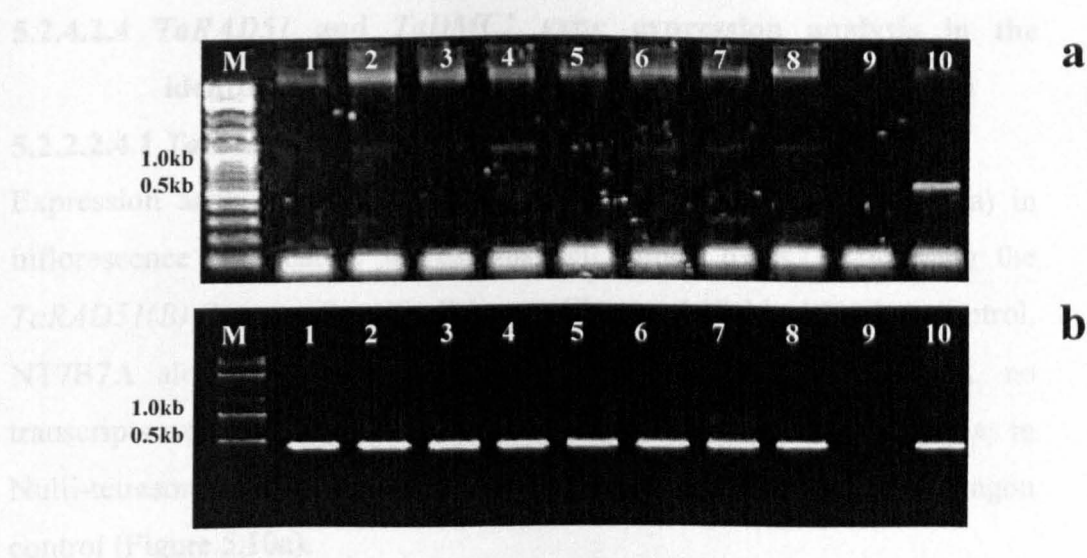


**Figure 5.7** PCR amplification of M<sub>3</sub> progeny of the *TaRAD51(B)* genome specific deletion line (line 81a) with the *TaRAD51(B)* Genome-Specific Primers (a), PCR amplification with *TaRAD51(A)* and (D) Genome-Specific Primers (b). 1-12 are M<sub>3</sub> progeny derived from line 81a M<sub>2</sub> parent; 13-water control and 14-Paragon control; the absence of PCR products in lanes 1 and 9 suggests low DNA quantity for PCR amplification; M, 2-log ladder.

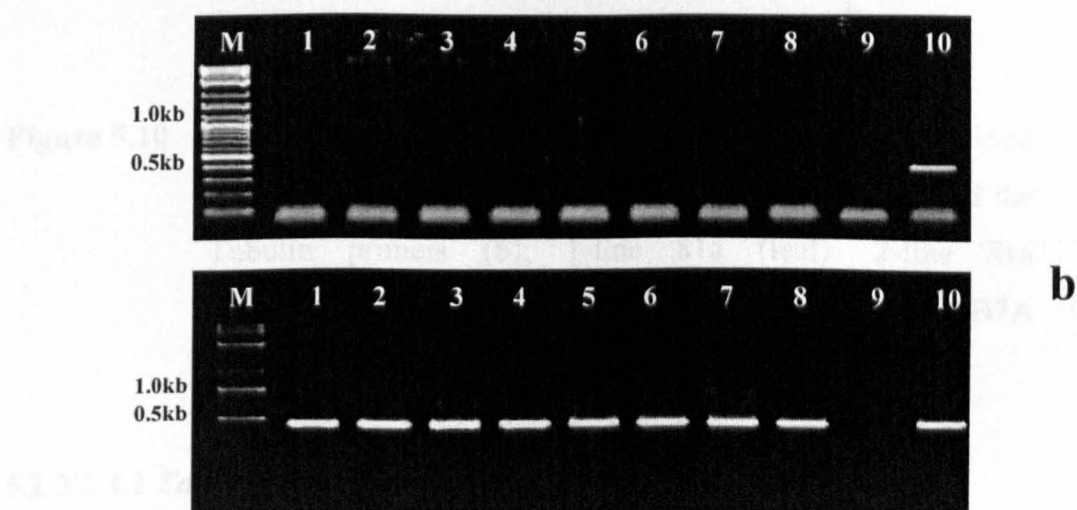
#### 5.2.4.2.3.2 Confirmation of the transmission of the deletion and Forward Genetic screening of the *TaDMC1* homoeologous gene mutants

Only the *TaDMC1(A)* and (D) homoeologous gene mutants identified by PCR survived (lines 176c and 260a) and no seed from previous generation (M<sub>1</sub>) was available to. Eight M<sub>3</sub> lines for each of M<sub>2</sub> parent (lines 176c and 260a) were planted. Before examining the deletion lines for Forward Genetics, the transmission of the deletion in the M<sub>3</sub> progeny was confirmed by PCR amplifications with *TaDMC1(A)* and (D) Genome-Specific Primers, as tests and as controls, depending upon the lines being examined (Table 5.2). As expected, the PCR amplifications indicated that the deletion is transmitted into the M<sub>3</sub> progeny for all the lines tested (Figure 5.8 and 5.9). Visible phenotypes, if there were any, were scored regularly during the growing season including key developmental time points for all the M<sub>3</sub> progeny in the glasshouse.





**Figure 5.8** PCR amplification of the M<sub>3</sub> progeny of the *TaDMC1(A)* Genome-Specific deletion line (line 176c) with *TaDMC1(A)* Genome-Specific Primers (a); PCR amplification with the *TaDMC1(D)* Genome-Specific Primers (b); 1-8 are M<sub>3</sub> progeny derived from *TaDMC1(A)* genome-specific M<sub>2</sub> parent (line 176c), 9-water control and 10-Paragon control; M, 2-log ladder.

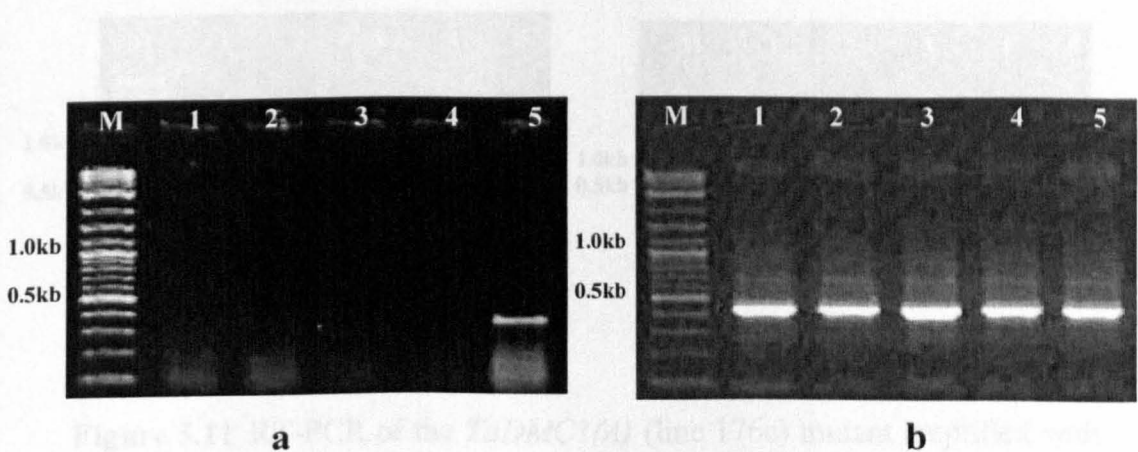


**Figure 5.9** PCR amplification of the M<sub>3</sub> progeny of the *TaDMC1(D)* Genome-Specific deletion line (line 260a) with the *TaDMC1(D)* Genome-Specific Primers (a); PCR amplification with *TaDMC1(A)* Genome-Specific Primers (b); 1-8 are M<sub>3</sub> progeny derived from *TaDMC1(D)* genome specific M<sub>2</sub> parent (line 176c), 9-water control and 10-Paragon control; M, 2-log ladder.

#### 5.2.4.2.4 *TaRAD51* and *TaDMC1* gene expression analysis in the identified Paragon deletion lines and unmutated Paragon

##### 5.2.2.2.4.1 *TaRAD51* expression analysis

Expression analysis for the identified *TaRAD51(B)* mutant (line 81a) in inflorescence and mature leaves was determined by RT-PCR using the *TaRAD51(B)* Genome-Specific Primers (Chapter 4, Table 4.2). As a control, NT7B7A along with unmutated Paragon were used. As expected, no transcripts were detected in the *TaRAD51(B)* mutant (line 81a) as well as in Nulli-tetrasomic 7B7A but transcripts were found in the unmutated Paragon control (Figure 5.10a).



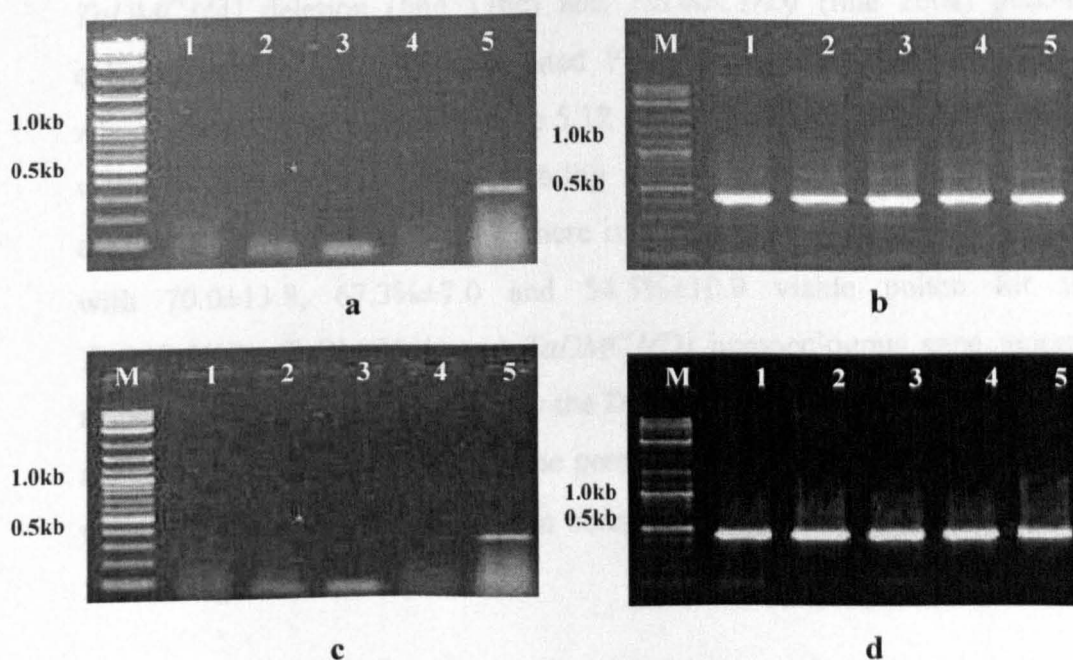
**Figure 5.10** RT-PCR of the *TaRAD51(B)* mutant (line 81a) amplified with the *TaRAD51(B)* Genome-Specific Primers (a) and the Tubulin primers (b); 1-line 81a (leaf), 2-line 81a (Inflorescence), 3-NT7B7A (leaf), 4-NT7B7A (inflorescence) and 5-Paragon (inflorescence).

##### 5.2.2.2.4.2 *TaDMC1* expression analysis

*TaDMC1* expression analysis by RT-PCR was determined for the identified *TaDMC1(A)* (line 176c) and *TaDMC1(D)* (line 260a) genome specific mutants along with unmutated Paragon using the *TaDMC1* Genome-Specific Primers (Chapter 4, Table 4.2). As a control, the same sets of primers were used to amplify unmutated Paragon. As expected, mRNA was not detected in the *TaDMC1(A)* mutant (line 176c) and *TaDMC1(D)* mutant (line 260a) with *TaDMC1(A)* and (D) Genome-Specific Primers,



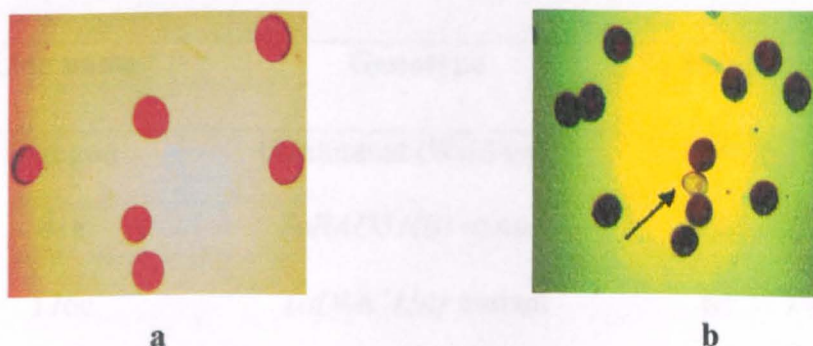
respectively, but expression was found in the unmutated Paragon control (Figure 5.11a and c).



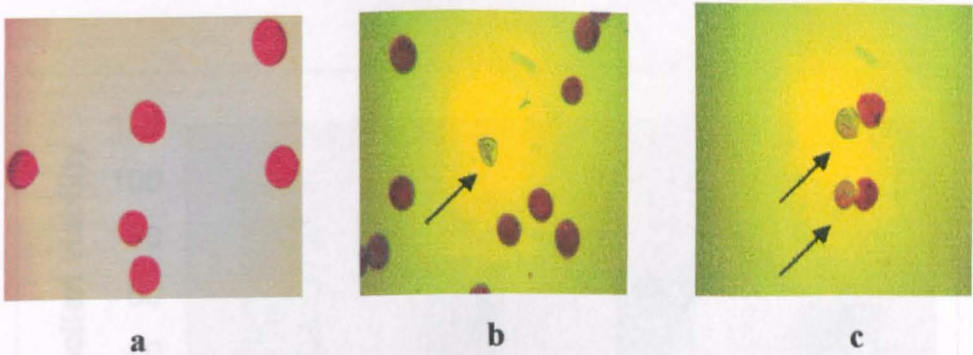
**Figure 5.11** RT-PCR of the *TaDMC1(A)* (line 176c) mutant amplified with the *TaDMC1(A)* Genome-Specific Primers (a) and Tubulin primers (b); 1-line 176c (leaf), 2-line 176c (inflorescence), 3-NT5A5D (leaf), 4-NT5A5D (inflorescence) and 5-Paragon (inflorescence); RT-PCR of the *TaDMC1(D)* (line 260a) mutant amplified with the *TaDMC1(D)* Genome-Specific Primers (c) and Tubulin primers (d); 1-line 260a (leaf), 2-line 260a (inflorescence), 3-NT5D5A (leaf), 4-NT5D5A (inflorescence) and 5-Paragon (inflorescence);

#### 5.2.4.2.5 Pollen viability assay in the identified Paragon deletion lines of *TaRAD51* and *TaDMC1* homoeologous genes and unmutated Paragon

The pollen viability assay of the *TaRAD51(B)* deletion (line 81a), *TaDMC1(A)* deletion (line 176c) and *TaDMC1(D)* (line 260a) genome deletion lines along with unmutated Paragon was determined using the Alexander staining method (Figure 5.12 and 5.13). The results of the pollen viability showed 99% pollen viability for the unmutated Paragon plants analyzed. For the deletion lines, there is a significant reduction in viability with  $70.0 \pm 13.8$ ,  $67.3\% \pm 7.0$  and  $54.5\% \pm 10.9$  viable pollen for the *TaRAD51(B)*, *TaDMC1(A)* and *TaDMC1(D)* homoeologous gene mutants respectively. Statistical analysis by the Duncan method revealed that there is a significant difference between the percentage of pollen viability of all the three homoeologous mutants when compared to unmutated Paragon (Table 5.5 and Figure 5.14).



**Figure 5.12** Pollen viability assay with Alexander's stain in unmutated Paragon plants (a) and line 81a (*TaRAD51(B)*) mutant plants (b). The arrow indicate the non-viable pollen.



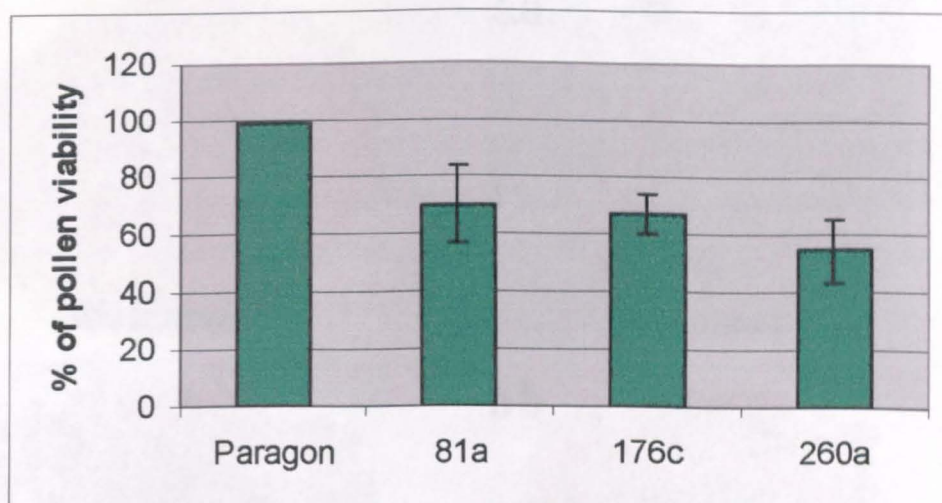
**Figure 5.13** Pollen viability assay with Alexander’s stain in unmutated Paragon plants (a) line 176c (*TaDMC1(A)*) mutant plants (b) and line 260a (*TaDMC1(D)*) mutant plants (c). The arrows indicate non-viable pollen.

**Table 5.5** Pollen viability percent in unmutated Paragon, *TaRAD51(B)* (line 81a), *TaDMC1(A)* (line 176c) and *TaDMC1(D)* (line 260a) mutant lines as determined by Alexander staining.

Line name	Genotype	% Pollen viability
Paragon	Unmutated (Wild-type)	99±0.25 <sup>b</sup>
81a	<i>TaRAD51(B)</i> mutant	70.0±13.8 <sup>a</sup>
176c	<i>TaDMC1(A)</i> mutant	67.3±7.0 <sup>a</sup>
260a	<i>TaDMC1(D)</i> mutant	54.5±10.9 <sup>a</sup>

Values are means of six replicates per treatment +s.e. Values for the same parameter labeled with the same letter are not significantly different (P≤0.05)

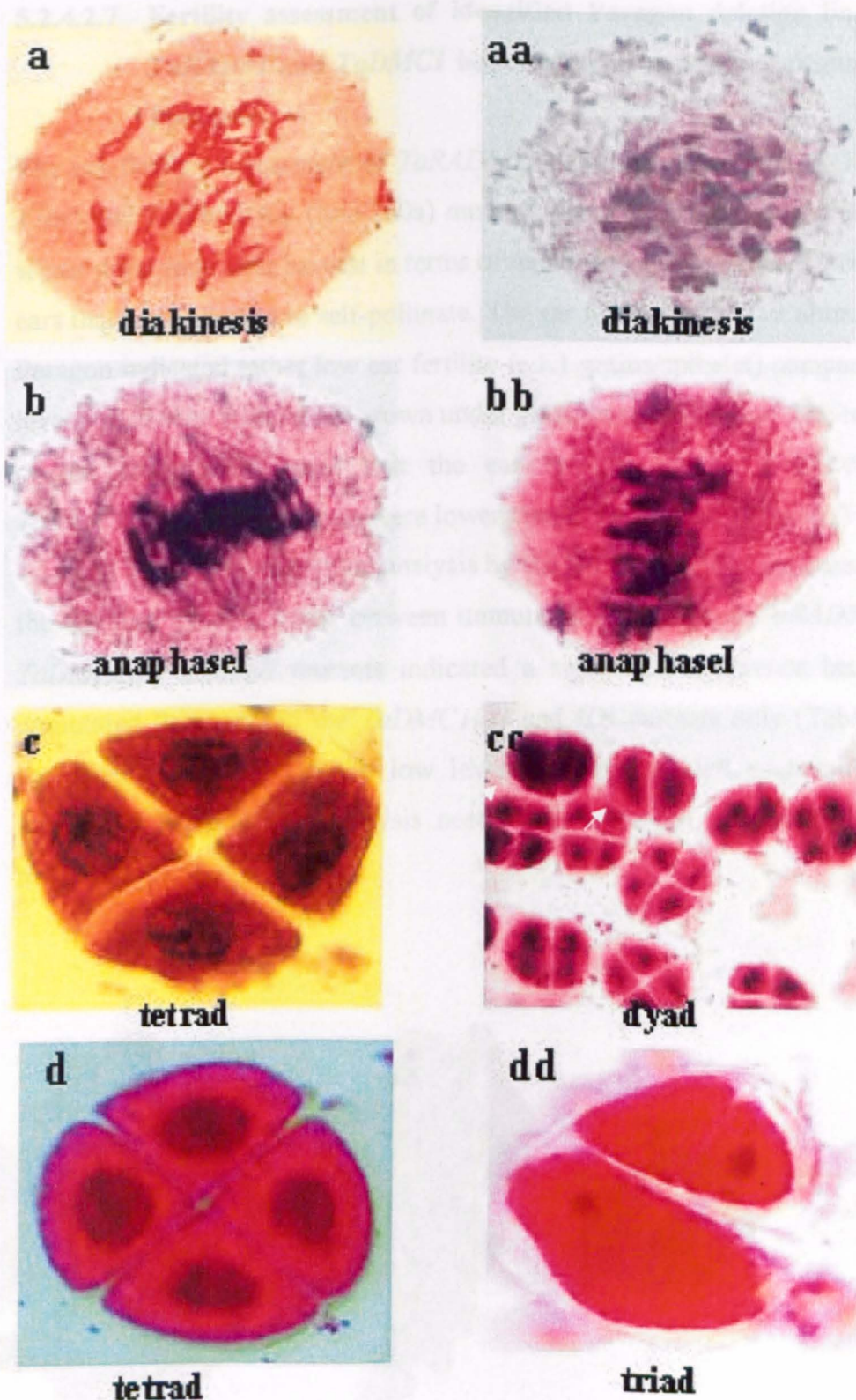




**Figure 5.14** A bar graph showing the percent of pollen viability in *TaRAD51(B)* (line 81a), *TaDMC1(A)* (line 176c) and *TaDMC1(D)* (line 260a) mutant lines compared with unmutated Paragon.

#### 5.2.4.2.6 Meiotic assessment of identified Paragon deletion lines of the *TaRAD51* and *TaDMC1* homoeologous genes and comparison with unmutated Paragon

Cytogenetic analysis of meiocytes in anthers was performed using Aceto-orcein staining of *TaRAD51(B)* (line 81a), *TaDMC1(A)* (line 176c) and *TaDMC1(D)* (line 260a) mutants along with unmutated Paragon. No major differences between the *TaRAD51(B)* (line 81a) mutant and the unmutated Paragon were observed at any stage of meiosis. The *TaRAD51(B)* mutant behaved normally in male meiosis and no abnormalities were detected indicating that the presence of non-viable pollen could be because of factors other than male meiosis. However slight abnormalities at diakinesis (Figure 5.15aa), early anaphase I (Figure 5.15bb) and the final tetrad stage (Figure 5.15cc and dd) were observed in some meiotic spreads of the *TaDMC1* mutants. No other abnormalities were detected in other stages of meiosis in these mutants. How far these abnormalities lead to a reduction in pollen viability needs further investigation.

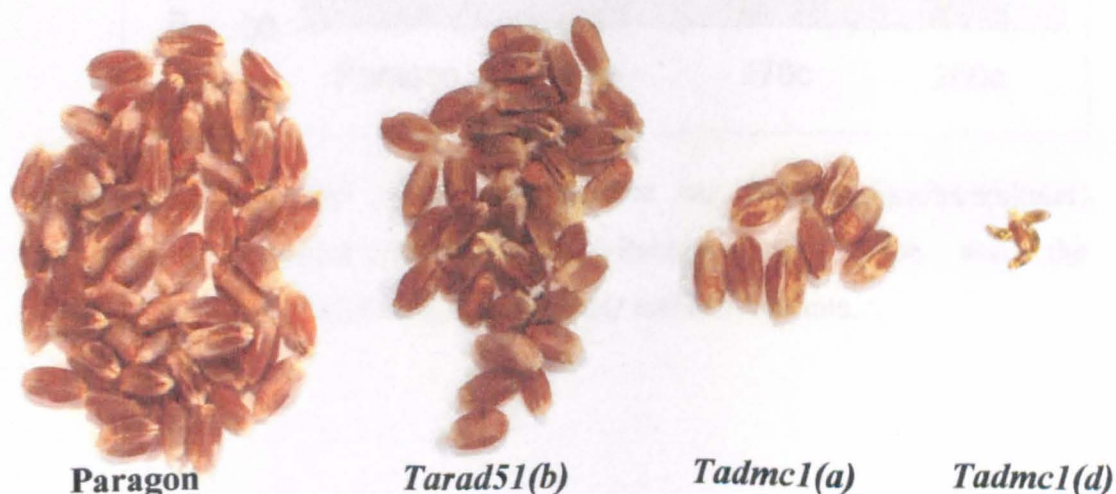


**Figure 5.15** Male meiosis I in unmutated Paragon wild type plant (a to d) compared with *TaDMC1(A)* and *TaDMC1(D)* mutant plants (aa to dd). The white arrow indicates dyads among tetrads in cc.



#### 5.2.4.2.7 Fertility assessment of identified Paragon deletion lines of *TaRAD51* and *TaDMC1* homoeologous genes and unmutated Paragon

The ear fertility assessment of *TaRAD51(B)* (line 81a), *TaDMC1(A)* (line 176c) and *TaDMC1(D)* (line 260a) mutants along with unmutated Paragon wheat were done after harvest in terms of number of grains/spikelet from six ears that were allowed to self-pollinate. The ear fertility results of unmutated Paragon indicated rather low ear fertility (~1.1 grains/spikelet) compared to levels often seen with wheat grown under glasshouse conditions. The results of the analysis indicated that the ear fertility of the *TaRAD51(B)*, *TaDMC1(A)* and *(D)* mutants were lower than in unmutated Paragon (Figure 5.16 and Table 5.6). Statistical analysis by the Duncan method indicated that the ear fertility comparing between unmutated Paragon and *TaRAD51(B)*, *TaDMC1(A)* and *(D)* mutants indicated a significant difference between unmutated Paragon and the *TaDMC1(A)* and *(D)* mutants only (Table 5.6 and Figure 5.17). Given the low levels of fertility seen even with the unmutated Paragon, this analysis needs to be repeated, preferably in the field.

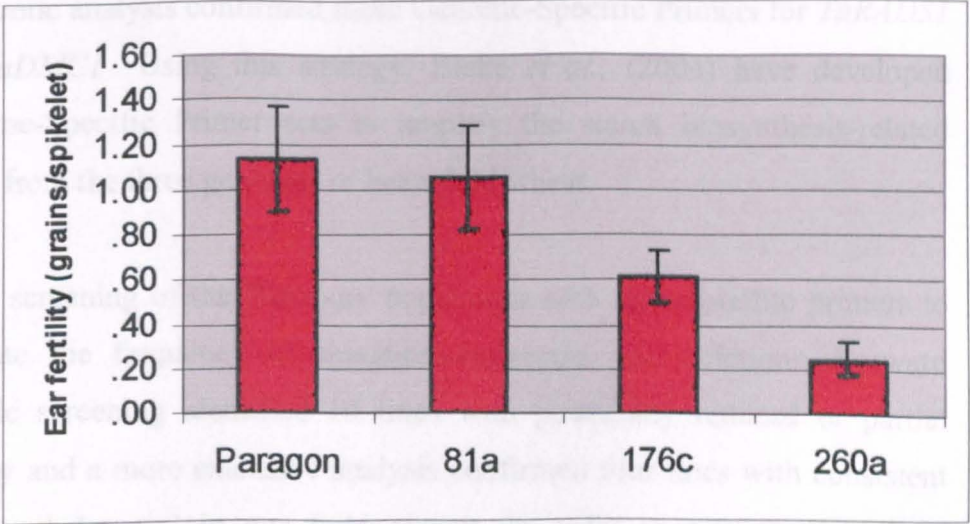


**Figure 5.16** Ear fertility (grains/spikelet) comparison between unmutated Paragon and the *TaRAD51(B)*, *TaDMC1(A)* and *TaDMC1(D)* mutants.

**Table 5.6** Ear fertility (grains/spikelet) values in unmutated Paragon, *TaRAD51(B)*, *TaDMC1(A)* and *(D)* mutants after harvest.

Line name	Genotype	Ear fertility (grains/spikelet)
Paragon	Wild type	1.1342±0.23 <sup>a</sup>
81a	<i>TaRAD51</i> B mutant	1.0570±0.23 <sup>a</sup>
176c	<i>TaDMC1</i> A mutant	0.6144±0.12 <sup>ab</sup>
260a	<i>TaDMC1</i> D mutant	0.2467±0.07 <sup>b</sup>

Values are means of six replicates per treatment +s.e. Values for the same parameter labeled with the same letter are not significantly different (P≤0.05)



**Figure 5.17** A bar graph showing the ear fertility (grains/spikelet) comparisons between Paragon wild type and the *TaRAD51(B)*, *TaDMC1(A)* and *(D)* mutants.



### 5.3 Discussion

Genome-specific primer sets were designed for the analysis of *TaRAD51* and *TaDMC1* genes for Forward and Reverse Genetics, as such primer sets would in effect allow hexaploid wheat to be treated as a diploid species with regard to DNA-based molecular analysis. In the absence of the genomic sequence of *TaRAD51* and *TaDMC1* for the design of Genome-Specific Primers within the introns (as the exons are nearly identical), the cDNA sequences of the *TaRAD51* and *TaDMC1* genes were aligned with the *OsRAD51A1* and the *OsDMC1A* cDNA sequences respectively, allowing the provisional identification of exon/intron boundaries and design of exon-anchored primers which amplify intronic regions. Variation within the intronic regions of these genes once amplified from genomic DNA allowed the design of putative Genome-Specific Primers. Subsequent Nulli-tetrasomic analysis confirmed these Genome-Specific Primers for *TaRAD51* and *TaDMC1*. Using this strategy, Blake *et al.*, (2004) have developed Genome-Specific Primer sets to amplify the starch biosynthesis-related genes from the three genomes of hexaploid wheat.

Initial screening of the Highbury population with microsatellite primers to estimate the frequency of mutation suggested few deletions. Forward Genetic screening identified 10 lines with potentially reduced or partial fertility and a more extensive analysis confirmed four lines with consistent grains/spikelet < 2 in one field season. In order to confirm that these mutants arise because of deletions in target genes, PCR amplification using *TaRAD51* and *TaDMC1* Genome-Specific Primers was performed. The results indicated that none of the putative lines showing reduced fertility contained deletions for these genes, which is not unexpected given that there are many potential causes of reduced fertility. Given the low levels of Genome-Specific deletions identified, the full Highbury population was not screened with the *TaRAD51* and *TaDMC1* Genome-Specific Primers. Reduced fertility lines have been stored for later analysis. The lack of deletions detected using microsatellite analysis might be explained on the basis that M<sub>2</sub> and successive generations of Highbury were produced by self-fertilization under single-seed descent, resulting in ~97% of the

mutations being homozygous at *M<sub>6</sub>*, while losing a quarter of all unfixed mutations at each generation. Even weak selection against the presence of deletions could lead to a reduction in their frequency over the generations.

In the second experiment, a gamma mutagenized population of Paragon wheat was used to characterize the *TaRAD51* and *TaDMC1* homoeologues. Initial screening of the population with *TaRAD51* and *TaDMC1* Genome-Specific Primers identified four and five lines with deletions for target genes (*TaRAD51* and *TaDMC1*, respectively), covering all six homoeologues. Unfortunately not all of the *M<sub>2</sub>* lines survived in the glasshouse, probably due to a very high mutational load and a potentially deleterious phenotype leading to reduced fertility (as seen in the remaining lines). Before subjecting the *M<sub>3</sub>* progeny of the homoeologous deletion lines of *TaRAD51* and *TaDMC1* to Forward Genetics, the presence of the deletion was confirmed using the *TaRAD51* and *TaDMC1* Genome-Specific Primers. The PCR analysis of the *M<sub>3</sub>* population indicated that the deletion was fixed and transmitted into the *M<sub>3</sub>* lines analyzed. The RT-PCR analysis of unmutated Paragon, *TaRAD51(B)*, *TaDMC1(A)* and *TaDMC1(D)* mutants revealed the deletion of the expected homoeologous genes in the *M<sub>3</sub>* progeny of the mutants which resulted in an absence of expression of the expected transcript. The unmutated Paragon wheat has all three homoeologous transcripts in all tissues analyzed for *TaRAD51* and *TaDMC1*.

The results of data analysis at various growth stages including key developmental time points indicated no significant differences in vegetative growth of the deletion lines compared with unmutated Paragon. This is also in accordance with the report in *atrads51* and *atdmc1* mutants in *Arabidopsis* (Couteau *et al.*, 1999 and Li *et al.*, 2004), *Zmrad51a1* and *Zmrad51a2* mutants in maize (Li *et al.*, 2007) and *OsDMC1*-RNAi line in rice (Deng and Wang, 2007) where the loss of function of either the *RAD51* gene or the *DMC1* gene has no visible effect on vegetative growth and so loss of a single copy of one of the homoeologues of *RAD51* and *DMC1* genes in wheat would also be predicted to have no clear effect on the vegetative growth of the plant. Given that wheat is a hexaploid, it is likely that the

other two homoeologues provide partial insulation against the effect of mutations in individual homoeologues (Stadler, 1929). So in order to elucidate the role of *RAD51* and *DMC1* homoeologous genes at various vegetative growth stages in hexaploid wheat, the availability of double and even triple deletions of *RAD51* and *DMC1* homoeologous genes is necessary.

The investigation of the effect of *TaRAD51(B)*, *TaDMC1(A)* and *TaDMC1(D)* deletions at the reproductive stages of the mutant using a pollen viability assays gave an initial indication of a significant difference in pollen viability in the mutant lines compared to unmutated Paragon. But the pollen viability results in this research are in sharp contrast to the reports published with other species. The pollen viability of the *TaRAD51(B)* mutant was found to be around 70% compared to 99% for the unmutated Paragon. By contrast the pollen viability of *atrads51-1* was found to be 0% (Li *et al.*, 2004). The pollen viability of *TaDMC1(A)* and *(D)* mutants were found to be around 67% and 55%, respectively. This is in contrast to the *atdmc1-1* mutant where the mutant shows just 1.5% pollen viability (Couteau *et al.*, 1999) and pollen in the RNAi knockout line of rice which showed less than 10% viability (Deng and Wang, 2007).

To correlate the reduction in pollen viability of the wheat mutants with abnormalities in meiosis - particularly male meiosis, cytogenetic analysis of male meiocytes was done. No meiotic abnormalities were observed in the *TaRAD51(B)* mutant indicating that the *TaRAD51(B)* homoeologue is not necessary for proper functioning of male meiosis in hexaploid wheat and that the deletion of the *TaRAD51(B)* gene is not necessarily the direct cause of the low pollen fertility observed here. This assumption is again in contrast with the *atrads51-1* mutant where *AtRAD51* gene is essential for normal male and female meiosis (Li *et al.*, 2004). It can be speculated that the other two *TaRAD51* homoeologues might be the important versions for male meiosis in wheat. However, a few meiotic abnormalities at diakinesis, prophaseI and the tetrad stage (frequent dyads) were observed for the *TaDMC1(A)* and *(D)* mutants. Similar but more drastic meiotic

abnormalities were observed in *atdmc1-1* mutant in which the homologous chromosomes at diakinesis stage were completely unpaired and ten univalents were found (five bivalents were seen in the wild-type plants) scattered throughout the cytoplasm (Couteau *et al.*, 1999).

Finally ear fertility analysis was carried out for the *TaRAD51(B)*, *TaDMC1(A)* and *TaDMC1(D)* mutant lines and compared with unmutated Paragon. No significant difference in ear fertility between *TaRAD51(B)* gene mutant (located on 7B chromosome) and unmutated Paragon was observed. Similar observations were made by Sears (1954) and found that Nullisomic of group 7 (nulli-7A, nulli-7B and nulli-7D) differ very little from euploid at the seedling and maturity stages and are distinguishable only by a slight reduction in vigor and height and by certain spike characters. The seed fertility and grain production of Nullisomic of 7B and 7D are nearly equal to euploid but is greatly reduced in Nullisomics of 7A due to pistilloidy. Endo and Gill (1996) has observed almost complete fertility as well with deletion lines of group 7 chromosomes. Both these reports confirm the current results and indicate that genes present on 7B chromosome are redundant (including *TaRAD51*) and partial or complete deletion of them doesn't have major effect at any stage of the plant development.

However, the ear fertility of *TaDMC1(A)* and *(D)* gene mutants (located on 5A and 5D respectively) was significantly different from unmutated Paragon and meiotic abnormalities appear to be partially responsible for the lower ear fertility of the *TaDMC1(A)* and *(D)* mutants. Sear (1954) observed that Nullisomics of group 5 have narrow leaves and slender culms, are late in maturity, spikes are reduced in size and have small glumes and seeds; they are female fertile and male sterile. Because critical genes for male fertility are located in those chromosomes (Sears and Sears 1978) it is not surprising to find very low levels of fertility or in some cases complete sterility in Nullisomic lines. Chromosomes 5D and 5A have an effect on premeiotic association opposite to that of their homoeologue 5B. It is assumed that the 5B gene was derived as an antimorphic mutation from an

association-promoting gene like those carried by 5D and 5A (Feldman, 1966). Endo and Gill (1996) observed that homozygotes for larger deletions in the long arms of chromosomes 5A, 5B, and 5D could not be obtained because of sterility. And as the size of the deletion increased in 5A chromosome, especially over 50% of the long arm, the seed set decreased to several grains per spike. Sears and Sears (1978) reported that the ditelosomic 5DS is male sterile. The above reports confirm the current results of the research. Because no complete sterility is observed with deletions in these homoeologues it can be speculated there might be other important genes on these chromosomes which are more important than *DMC1* gene for ear fertility. The difference in the levels of fertility between the *TaDMC1(A)* and *(D)* deletion lines may be because of their relative importance of the homoeologues in meiosis.

In summary, results from evaluation of the wheat homoeologous deletions for *TaRAD51(B)*, *TaDMC1(A)* and *TaDMC1(D)* and *Arabidopsis rad51* and *dmc1* clearly demonstrated that both the genes are dispensable for vegetative growth stages of the plant. However, differences were observed in the reproductive stages of these gene mutants. While *atrads51* mutants are male and female sterile, the *TaRAD51(B)* mutant is male fertile (although with a significantly reduced pollen viability) and does not display a significantly reduced seed set. This result is similar in some ways and contrasts in others with the observation in *Zmrad51A1* and *Zmrad51A2*, that homozygous single mutants of either gene develop normally, are fully male fertile, and do not display reduced seed set (Li *et al.*, 2007). This result is also to the observation of seed fertility in Nullisomic of 7B chromosomes. From the results of *Arabidopsis* and maize mutants and also from Nullisomic data, it can be concluded that in wheat *TaRAD51(B)* appear to be functionally redundant *in planta*. However future experiments involving the other *TaRAD51* homoeologue mutants (*A* and *D*) may give a clearer picture. In the case of the *atdmc1* mutant, even though meiotic abnormalities were found it does not lead to complete sterility. Similar results (albeit less dramatic ones) were observed for *TaDMC1(A)* and *TaDMC1(D)* mutants. This argues that the function of *DMC1* is conserved among plants and

observed abnormalities in meiosis do not lead to cell death as in the case of yeast and mouse. From these initial results it can be concluded there is likely to be a dosage effect of *TaDMC1* homoeologues on the observed phenotypes. However future experiments involving the *TaDMC1(B)* homoeologue mutant and combinations of these genes may give a clearer picture. Perhaps an unexpected result for the *TaDMC1(A)* and *TaDMC1(D)* mutant lines is that the deletion of a single copy of these should have a phenotypic effect on meiosis (assuming that these mutations are causing the mild disruption of meiosis) which argues that the buffering capacity is limited for these homoeologues. Finally the results of deletions of *TaRAD51* and *TaDMC1* needs to be considered with caution because apart from the confirmed deletions there can be other deletions in those lines and the observed phenotype may not be true indication of the deleted gene. So it is important to backcross the deletion mutant with original parent line from which it is derived to remove any additional deletions. The deletion lines for target genes (*TaRAD51* and *TaDMC1*) in this study have only been studied for one generation and so to relate the observed phenotype to the deletions in target genes only and to remove any additional deletions, future research should aim to backcross the mutants with parent.

## **CHAPTER 6: FUNCTIONAL CHARACTERISATION OF THE *TaRAD51* AND *TaDMC1* GENES IN *Arabidopsis thaliana***

### **6.1 Introduction**

A variety of approaches such as mutagenesis, RNAi etc., have been used to disrupt the function of genes in plants. Insertional mutagenesis has been extensively used, particularly in *Arabidopsis*. It is based on the insertion of unrelated DNA into the gene of interest. In *Arabidopsis*, insertional mutagenesis involves the use of transposable elements (Sundaresan *et al.*, 1995; Martienssen, 1998) or the T-DNA of *Agrobacterium tumefaciens* (Azpiroz-Leehan and Feldmann, 1997). The introduced foreign DNA not only disrupts the production of viable transcripts of the target gene but also acts as a marker for identification of the mutated locus in subsequent generations. In *Arabidopsis*, insertional mutagenesis has been utilized to produce large numbers of mutants. The success of this approach is partly due to small intron size and high density of coding sequence – a reflection of the small genome size of *Arabidopsis*. The T-DNA insertion is in the order of 5 to 25 kb in length and generally produces a loss of the gene function in *Arabidopsis*, although there are numerous examples of the T-DNA being excised at the transcript level to give functional expression of the gene (Matzke *et al.*, 1994). Mutations that are homozygous lethals can also be maintained within lines in the form of heterozygous plants.

Insertional mutagenesis lines for many genes have been generated in *Arabidopsis* by various research groups throughout the world. The major groups of lines generated for research and the research groups involved are SALK lines (SALK Institute Genomic Analysis Laboratory, USA) (Alonso *et al.*, 2003), GABI-Kat lines (Institute of Genome Research, Bielefeld



University, Germany) (Rosso *et al.*, 2003) and SAIL lines (Syngenta Arabidopsis Insertion Library, USA) (Sessions *et al.*, 2002). These large collections of T-DNA insertion lines are curated and are provided by the NASC to Europe and by the Arabidopsis Stock Center at Ohio State University for the US.

In Arabidopsis both the characterised *RAD51* and *DMC1* gene mutants are generated by T-DNA insertional mutagenesis and have been genetically well characterized (Couteau *et al.*, 1999; Li *et al.*, 2004). Insertional mutagenesis into both the *AtRAD51* and *AtDMC1* genes produces a loss of the functional allele. Homozygous *Atrad51* plants show no apparent abnormal phenotype during the vegetative phase, but pollen from the homozygous *atrads51* plants is non-viable and seed set does not occur. Similarly, homozygous *atdmc1* plants developed normally in the vegetative stage, but show strongly reduced fertility during the reproductive phase.

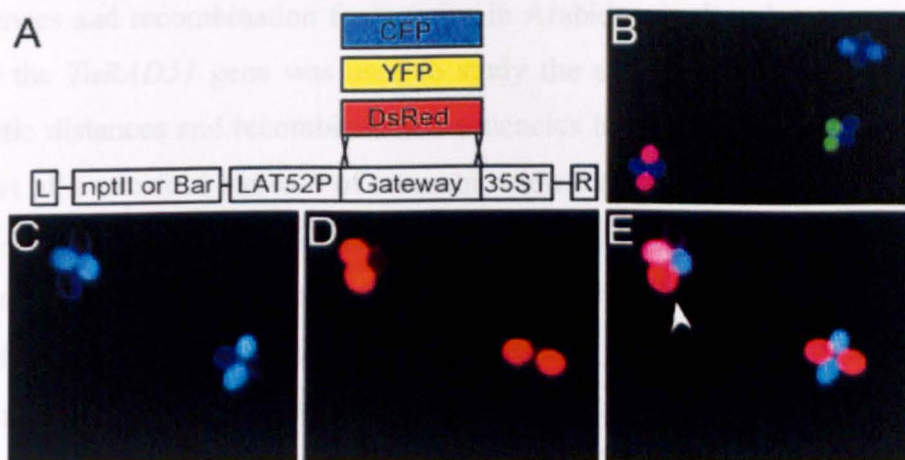
Phylogenetic analysis based on amino acid sequences show significant homologies between the Arabidopsis (a dicot species) and the wheat (a monocot species) *RAD51* and *DMC1* genes (Chapter 4). So the functions of *TaRAD51* and *TaDMC1* could be investigated by complementing the Arabidopsis *rad51* and *dmc1* mutant phenotypes by over-expression of the corresponding wheat allele. Any complementation of Arabidopsis mutant phenotype with *TaRAD51* and *TaDMC1* genes would be a strong evidence for functional relatedness. Delhaize *et al.*, (1999) showed that yeast *cho1* mutants that lacked PSS (phosphatidylserine synthase) activity can be complemented by the expression of *TaPSS1* gene.

Li *et al.* (2004) showed that the *Atrads51* mutant can be complemented by the introduction of a *CAMV35S::MYC::AtRAD51* construct into heterozygous *AtRAD51/Atrads51* plants. This completely restored fertility in the homozygous *atrads51* mutants. Couteau *et al.*, (1999) also showed that the *Atdmc1* mutant

can be complemented by the introduction of an 11-kb DNA fragment covering the full sequence of *AtDMC1* (Columbia ecotype) used into heterozygous (*AtDMC1/Atdmc1*) plants. Both these reports and also numerous reports from other genes in *Arabidopsis* (Mysore *et al.*, 2000) indicate that the complementation method can successfully be utilized to functionally complement mutant phenotype by the wild-type version of the respective gene, derived from the same species. However, the use of these T-DNA mutants for complement studies involving genes from unrelated species has not been reported to date. In this current work, the (*D*) homoeologue of the *TaRAD51* and *TaDMC1* genes were cloned into over-expression constructs to test for potential complementation in *Arabidopsis rad51* and *dmc1* T-DNA lines respectively.

Francis *et al.*, (2007) reported a fluorescence-based tetrad analysis system for detecting meiotic recombination events such as crossover interference, gene conversion and evaluating cross-over frequencies in *Arabidopsis*. This analysis system is based on the *qrt* mutant background of *Arabidopsis* which produces tetrads in which the four meiotic products are held together by undegraded callose, thus allowing all four products of a single meiotic event to be studied relative to one another. This is coupled with a series of transgenic marker genes that express different coloured fluorescent proteins in pollen (Stewart, 2006). Each of the transgenic marker genes encodes one of the three fluorescent proteins (red, cyan and yellow) under the control of the post meiotic pollen-specific promoter (LAT52) (Figure 6.1). By transforming large numbers of plants with either a red, cyan or yellow construct, a series of lines containing random insertions into the five chromosomes of *Arabidopsis* were generated. Different colour markers can be combined through crossing or re-transformation to produce lines that carry linked different colour marker constructs. Genetic intervals can be visually assayed for recombination frequency by scoring for non-parental inheritance of linked markers. The *qrt* background allows this to be taken a step further, with all four ordered meiotic

products scorable. This system can thus be used to study the effect of over-expression of the wheat meiotic recombination genes (*TaRAD51* and *TaDMC1*) on genetic intervals and recombination frequencies in Arabidopsis.



**Figure 6.1** Fluorescence-based tetrad system for analyzing meiotic recombination events in Arabidopsis. An *Agrobacterium* T-DNA construct containing eCFP, eYFP and DsRED driven by LAT52 promoter (A);  $T_1$  plants expressing the fluorescent protein in their pollen (B); Crossing lines with differently colored transgenes on the same chromosome (C, D and E) enables detection of crossing over events in the interval between the transgenes (C and D with E, a merged image of C and D with arrow indicating recombinant tetrad) (Source: Francis *et al.*, 2007).

In the current chapter, molecular and phenotypic characterization of the Arabidopsis T-DNA insertion lines for the *RAD51* and *DMC1* genes was undertaken to confirm the reported results of Li *et al.*, (2004) and Couteau *et al.*, (1999). Complementing constructs were prepared with one of the wheat *RAD51* and *DMC1* homoeologues driven by a *CaMV35S* promoter and transformed into heterozygous mutant lines (*AtRAD51/AtRAD51* and

*AtDMC1/Atdmc1*). Functional complementation by over-expression of the wheat strand exchange genes in Arabidopsis was analyzed by phenotypic assessment and expression analysis by RT-PCR. Finally the fluorescence-based tetrad analysis system was provisionally used to examine whether there was any effect of over-expression of the *TaRAD51* and *TaDMC1* genes on genetic distances and recombination frequencies in Arabidopsis. For the current study, only the *TaRAD51* gene was used to study the effects of over-expression on genetic distances and recombination frequencies in Arabidopsis because of the report of a two-fold increase of recombination frequency upon overexpression of *RAD51* in Arabidopsis (Betzner and Stefan *et al.*, WO/2002/008432). But in the current work, overexpression was tested with *TaRAD51(D)* instead of *AtRAD51* version. This was initially undertaken without recovery of the *qrt* background. Construction and analysis of lines over-expressing *TaDMC1(D)* in a fluorescent tester background for recombination is underway.

## **6.2 Results**

### **6.2.1 Comparative amino acid sequence analysis between the *AtRAD51* and the *TaRAD51* cDNA homoeologues**

Comparative amino acid sequence analysis of *TaRAD51* (*cTaRAD51*-7A, *cTaRAD51*-7B and *cTaRAD51*-7D; FJ594479, FJ594480 and FJ5944781) homoeologous genes against *AtRAD51* (NM\_122092) was performed and the analysis revealed a high level of sequence conservation between the three cDNA homoeologues of *TaRAD51* and *AtRAD51* cDNA sequence, except for the first 30 amino acids at 5'- end of the sequences (Figure 6.2). As there is a high level of sequence conservation between the three cDNA homoeologues, for the current study only one cDNA homoeologues (*TaRAD51(D)*) was used to test for functional complementation of the *Atrad51* T-DNA insertional mutant and also for studying recombination frequencies in the fluorescent tetrad lines of Arabidopsis by over-expression.



```

TaRAD51A MSSAAHQKAAAAAPVEEEAGEHGFPIEHLQASGIAAVDVKKLRDAGLCTVESVAYSP 60
TaRAD51B MSSAAHQKAAAAAPVEEEAGEHGFPIEHLQASGIAAVDVKKLRDAGLCTVESVAYSP 60
TaRAD51D MSSAAHQKAAAAAPVEEEAGEHGFPIEHLQASGIAAVDVKKLRDAGLCTVESVAYSP 59
ATRAD51 MTTHGQRNRNNAVQQQDDDET-QHGFPPVEQLQAAGIASVDVKKLRDAGLCTVEGVAYTP 59
*::: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*: :*
TaRAD51A RKDLLQIKGISAKVDKIIAAASKLVPLGFTSATQLHAQRLEIIQVTTGSRDLKILEGC 120
TaRAD51B RKDLLQIKGISAKVDKIIAAASKLVPLGFTSATQLHAQRLEIIQVTTGSRDLKILEGC 120
TaRAD51D RKDLLQIKGISAKVDKIIAAASKLVPLGFTSATQLHAQRLEIIQVTTGSRDLKILEGC 119
ATRAD51 RKDLLQIKGISDAKVDKIIAAASKLVPLGFTSASQLHAQRQEIIQITSCSRDLKILEGC 119
*****:*****:*****:*****:*****:*****:*****:*****:*****:*****:*****
TaRAD51A IETGSITELYGEFRSGKTQLCHTLCTCQLPLDQGGGEGKALYIDAEGTFRPQRLQLIAD 180
TaRAD51B IETGSITELYGEFRSGKTQLCHTLCTCQLPLDQGGGEGKALYIDAEGTFRPQRLQLIAD 180
TaRAD51D IETGSITELYGEFRSGKTQLCHTLCTCQLPLDQGGGEGKALYIDAEGTFRPQRLQLIAD 179
ATRAD51 IETGSITELYGEFRSGKTQLCHTLCTCQLPHDQGGGEGKALYIDAEGTFRPQRLQLIAD 179
*****:*****:*****:*****:*****:*****:*****:*****:*****:*****
TaRAD51A RFGLNGADVLENVAYARAYNTDQSRLLLEAASMMHVEIRFALMVIDSATALYRTDFSGRG 240
TaRAD51B RFGLNGADVLENVAYARAYNTDQSRLLLEAASMMHVEIRFALMVIDSATALYRTDFSGRG 240
TaRAD51D RFGLNGADVLENVAYARAYNTDQSRLLLEAASMMHVEIRFALMVIDSATALYRTDFSGRG 239
ATRAD51 RFGLNGADVLENVAYARAYNTDQSRLLLEAASMMHVEIRFALLIVDSATALYRTDFSGRG 239
*****:*****:*****:*****:*****:*****:*****:*****:*****:*****
TaRAD51A ELSARQMHLAKFLRSLQKLADFCVAVVISNQVVAQVDGCGAMFAGPQIKPIGCNIMAHAS 300
TaRAD51B ELSARQMHLAKFLRSLQKLADFCVAVVISNQVVAQVDGCGAMFAGPQIKPIGCNIMAHAS 300
TaRAD51D ELSARQMHLAKFLRSLQKLADFCVAVVISNQVVAQVDGCGAMFAGPQIKPIGCNIMAHAS 299
ATRAD51 ELSARQMHLAKFLRSLQKLADFCVAVVITNQVVAQVDGCSALFAGPQFKPIGCNIMAHAT 299
*****:*****:*****:*****:*****:*****:*****:*****:*****:*****
TaRAD51A TTRLYLKKGRAERICKVVSPPCLAEAEARFQISPEGVTDVKD 343
TaRAD51B TTRLYLKKGRAERICKVVSPPCLAEAEARFQISPEGVTDVKD 343
TaRAD51D TTRLYLKKGRAERICKVVSPPCLAEAEARFQISPEGVTDVKD 342
ATRAD51 TTRLALPKGRAERICKVISSPCLPEAEARFQISTEGVTDCKD 342
**** *****:*****:*****:*****:***** **

```

**Figure 6.2** Comparison of the amino acid sequences of *TaRAD51* cDNA homoeologues with *AtRAD51* cDNA sequence. The residues in yellow indicate conserved amino acids, the residues in red indicate amino acids that are different between the three *TaRAD51* cDNA homoeologues and residues in blue indicate amino acids that are different between *TaRAD51* and *AtRAD51*. Numbers on the right indicate positions of the amino acid residues with the predicted protein sequence.

### 6.2.2 Comparative amino acid sequence analysis between the *AtDMC1* and *TaDMC1* cDNA homoeologues

Comparative amino acid sequence analysis of *TaDMC1* (*cTaDMC1*-5A, *cTaDMC1*-5B and *cTaDMC1*-5D; FJ594482, FJ594483 and FJ594484) homoeologues against *AtDMC1* (NM\_113188) was performed and revealed a high level of sequence conservation between the three cDNA homoeologues of



amino acids at 5'-end of the sequences. As there is a high level of sequence conservation between the three cDNA homoeologues, for the current study only one cDNA homoeologue (*TaDMC1(D)*) was used to test for the functional complementation of the *Atdmc1* T-DNA insertional mutant by over-expression.

Comparative amino acid analysis of *TaDMC1* and *AtDMC1* revealed that both the structural motifs and catalytic domain were conserved between both

TaDMC1-5A	MAPSKQYDEGGQLQLMEADRVVEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMM	60
TaDMC1-5D	MAPSKQYDEGGQLQLMEADRVVEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMM	60
TaDMC1-5B	MAPSKQYDEGGQLQLMEADRVVEEEECFESIDKLISQGINSGDVKKLQDAGIYTCNGLMM	60
ATDMC1	MMASLKAEETSQMLVEREENDEDEDLFEMIDKLI AQGINAGDVKKLQEA GIHTCNGLMM	60
* . . . : * . : * : * . : * : * * * * * : * * * * * : * * * * * : * * * * *		
TaDMC1-5A	HTKKS LTG I K G L S E A K V D K I C E A A E K L L S Q G F M T G S D L L I K R K S V V R I T T G S Q A L D E L L G	120
TaDMC1-5D	HTKKS LTG I K G L S E A K V D K I C E A A E K L L S Q G F M T G S D L L I K R K S V V R I T T G S Q A L D E L L G	120
TaDMC1-5B	HTKKS LTG I K G L S E A K V D K I C E A A E K L L S Q G F M T G S D L L I K R K S V V R I T T G S Q A L D E L L G	120
ATDMC1	HTKKNLTG I K G L S E A K V D K I C E A A E K I V N F G Y M T G S D A L I K R K S V V K I T T G C Q A L D D L L G	120
**** . *		
TaDMC1-5A	CGIETLCIT E A F G E F R S G K T Q L A H T L C V S T Q L P L H M H G C G N K R V A Y I D T E G T F R P E R I V P I	180
TaDMC1-5D	CGIETLCIT E A F G E F R S G K T Q L A H T L C V S T Q L P L H M H G C G N K R V A Y I D T E G T F R P E R I V P I	180
TaDMC1-5B	CGIETLCIT E A F G E F R S G K T Q L A H T L C V S T Q L P L H M H G C G N K R V A Y I D T E G T F R P E R I V P I	180
ATDMC1	CGIETSAIT E A F G E F R S G K T Q L A H T L C V T T Q L P T N M K G C G N K R V A Y I D T E G T F R P D R I V P I	180
**** . *		
TaDMC1-5A	AERFGMDANAVLDNIIYARAYTYEHQYNLLCLAAKMAE E P F R L L I V D S V I A L F R W D F S C	240
TaDMC1-5D	AERFGMDANAVLDNIIYARAYTYEHQYNLLCLAAKMAE E P F R L L I V D S V I A L F R W D F S C	240
TaDMC1-5B	AERFGMDANAVLDNIIYARAYTYEHQYNLLCLAAKMAE E P F R L L I V D S V I A L F R W D F S C	240
ATDMC1	AERFGMDPGAVLDNIIYARAYTYEHQYNLLCLAAKMS E E P F R I L I V D S I I A L F R W D F T C	240
***** . *		
TaDMC1-5A	RGELAERQQKLAQMLSRLTKIAEEFNVA V Y I T N Q V I A D P G C G M F I T D P K K P A C G H V L A H A	300
TaDMC1-5D	RGELAERQQKLAQMLSRLTKIAEEFNVA V Y I T N Q V I A D P G C G M F I T D P K K P A C G H V L A H A	300
TaDMC1-5B	RGELAERQQKLAQMLSRLTKIAEEFNVA V Y I T N Q V I A D P G C G M F I T D P K K P A C G H V L A H A	300
ATDMC1	RGELADRQQKLAQMLSRLTKIAEEFNVA V Y M T N Q V I A D P G C G M F I S D P K K P A C G H V L A H A	300
***** . *		
TaDMC1-5A	ATIRLM L R K S K G E Q R V C K I F D A P N L P E G E A V F Q I T T G C L M D V K D	344
TaDMC1-5D	ATIRLM L R K S K G E Q R V C K I F D A P N L P E G E A V F Q I T T G C L M D V K D	344
TaDMC1-5B	ATIRLM L R K S K G E Q R V C K I F D A P N L P E G E A V F Q I T T G C L M D V K D	344
ATDMC1	ATIRLLFRKSGDTRVCKVYDAPNLAAEAAS F Q I T Q G G I A D A K D	344
* * * * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : *		

**Figure 6.3** Comparison of amino acid sequences of *TaDMC1* homoeologues with *AtDMC1*. The residues in yellow indicate conserved amino acids, the residues in red indicate amino acids that are different between the three *TaDMC1* cDNA homoeologues and residues in blue indicate amino acids that are different between *TaDMC1* and *AtDMC1*. Numbers on the right indicate positions of the amino acid residues.



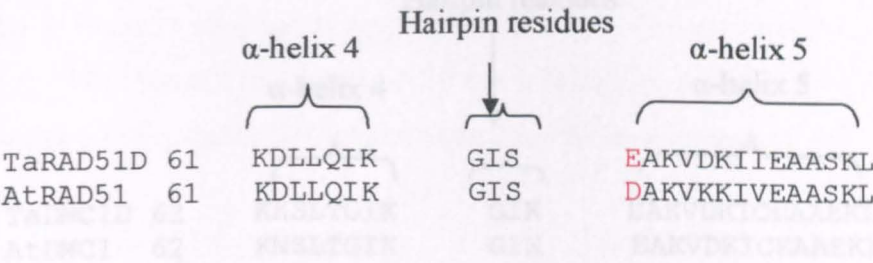
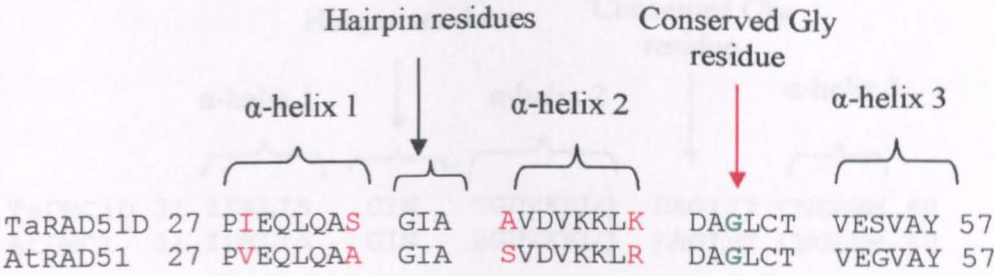
### **6.2.3 Comparative sequence analysis of the *TaRAD51(D)* and *AtRAD51* amino acids**

Comparative amino acid analysis of *TaRAD51(D)* against *AtRAD51* revealed the structural motifs and catalytic domains were spaced identically across both of these sequences even though amino acid differences were found in many of the motifs. The Walker A and Walker B motifs were also identically spaced (Figure 6.4).

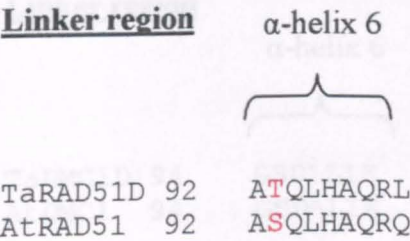
### **6.2.4 Comparative sequence analysis of the *TaDMC1(D)* and *AtDMC1* amino acids**

Amino acid sequence analysis of *TaDMC1(D)* against *AtDMC1* revealed very high levels of similarity to one another for all the domains analyzed. The structural motifs and catalytic domains were spaced identically across both of these sequences even though amino acid differences were found in many of the motifs between the sequences. The Walker A and Walker B motifs were also identically spaced (Figure 6.5).

**N-terminal domain**



**Linker region**



**C-terminal domain**

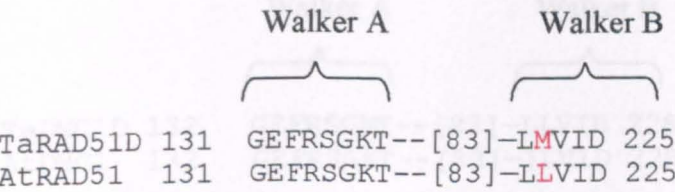
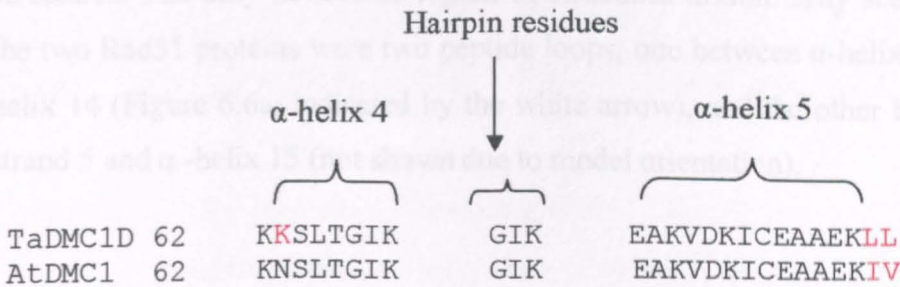
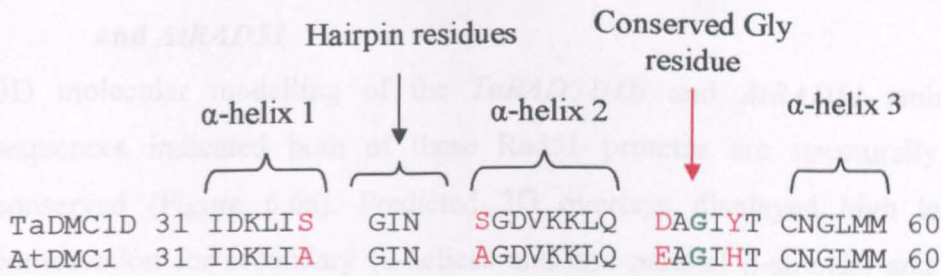


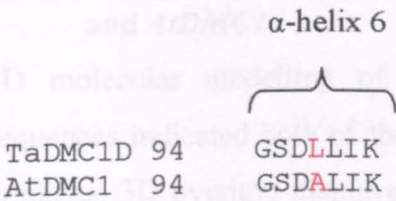
Figure 6.4 Comparative amino acid sequence analysis of *TaRAD51(D)*

**Figure 6.4** Comparative amino acid sequence analysis of *TaRAD51(D)* against *AtRAD51*. The amino acid residues that are different between the two genes are indicated in red. The conserved polar non-charged glycine residue (green) is indicated with a red arrow

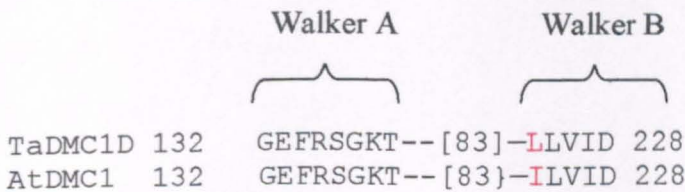
**N-terminal domain**



**Linker region**



**C-terminal domain**



**Figure 6.5** Comparative amino acid sequence analysis of *TaDMC1(D)* against *AtDMC1*. The amino acid residues that are different between the two genes are indicated in red. The conserved polar non-charged glycine residue (green) is indicated with a red arrow

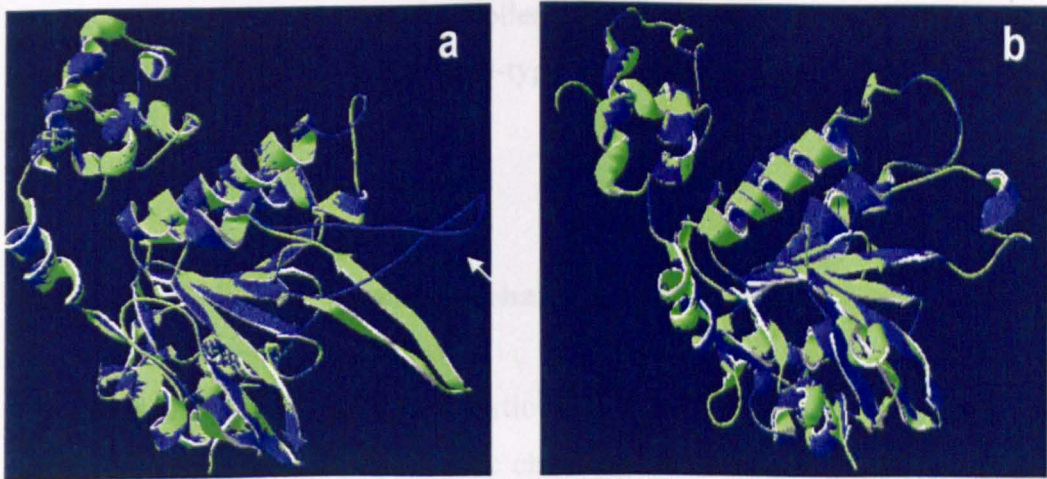
### **6.2.5 Structural conservation of the protein sequences of the *TaRAD51* and *AtRAD51***

3D molecular modelling of the *TaRAD51(D)* and *AtRAD51* amino acid sequences indicated both of these Rad51 proteins are structurally highly conserved (Figure 6.6a). Predicted 3D overlays displayed high levels of conservation for secondary ( $\alpha$ -helices and anti-parallel  $\beta$ -strands) and tertiary structures. The only noticeable region of structural dissimilarity seen between the two Rad51 proteins were two peptide loops; one between  $\alpha$ -helix 13 and  $\alpha$ -helix 14 (Figure 6.6a, indicated by the white arrow), and the other between  $\beta$ -strand 5 and  $\alpha$ -helix 15 (not shown due to model orientation).

### **6.2.6 Structural conservation of the protein sequences of the *TaDMC1* and *AtDMC1***

3D molecular modelling of the *TaDMC1(D)* and *AtDMC1* amino acid sequences indicated both of these Dmc1 proteins are conserved (Figure 6.6b). Predicted 3D overlays displayed high levels of conservation for secondary ( $\alpha$ -helices and anti-parallel  $\beta$ -strands) and tertiary structures.





**Figure 6.6** 3D predictive modeling of TaRad51(D) and AtRad51 superimposed on to each other (a); The 3D structure of TaRad51(D) is represented in green, and AtRad51 is represented in blue; The white arrow shows one of the peptide loop (AtRad51) that show disparity between the two proteins. The other peptide loop is not shown because of the images presented; 3D predictive modeling of TaDmc1(D) and AtDmc1 superimposed on to each other (b); The 3D structure of TaDmc1(D) is represented in green and AtDmc1 is represented in blue.

### 6.2.7 Molecular and phenotypic characterization of the *Atrad51* and *Atdmc1* T-DNA insertional lines

*Atrad51* and *Atdmc1* T-DNA insertional seed was obtained from Dr. Marie-Pascale Doutriaux (University Paris-Sud, France). To be able to use these lines for complementation experiments involving the *TaRAD51* and *TaDMC1* genes, these insertional mutants first underwent molecular and phenotypic characterization to confirm the expected genotype and phenotype were consistent with the reported results of Couteau *et al.*, (1999) and Li *et al.*, (2004). While phenotypic assessment of the plants was undertaken for

vegetative growth and also for pollen viability, genotyping of plants was performed by PCR using the wild-type and mutant allele specific primers (Table 2.5).

#### **6.2.7.1 Molecular and phenotypic characterization of *Arabidopsis rad51* T-DNA insertional mutants**

The *Arabidopsis rad51* T-DNA insertional mutant seed was obtained from Dr. Marie-Pascale Doutriaux. Systematic characterization of plants derived from a single *AtRAD51/Atrad51* plant was performed using PCR with primers specific for either the mutant or the wild-type *AtRAD51* alleles (Table 2.5) to define their genotypes. Systematic analysis of 25 plants derived from a single self-fertilized *AtRAD51/Atrad51* showed 7 of them to have just the wild-type alleles, 10 of them to have both alleles (indicating a heterozygote) and 8 of them to have the mutant allele to produce an approximate 1:2:1 ratios of the wild-type/ heterozygous/homozygous plants ( $\chi^2 = 0.749$ ;  $P=0.687$ ) (Figure 6.7).

Phenotypic characterization of the progeny derived from a self-fertilized heterozygous *AtRAD51/Atrad51* plant was also undertaken. Homozygous plants with the *atrad51/atrad51* genotype showed normal vegetative development as did the other two genotypes, with similar rates of growth and similar numbers of rosette and cauline leaves, but the double mutant plants were completely sterile and failed to set any seeds. The phenotypic ratio of the 25 plants gave an approximately 3:1 ratio in terms of fertile to sterility (17 fertile: 8 sterile;  $\chi^2 = 0.653$ , 1df = n.s.). The anthers of wild-type and heterozygous plants produced normal pollen whereas in the anthers of the homozygous *atrad51/atrad51* plants no viable pollens grains were observed and the grains were variable in shape and much smaller than for the wild-type and heterozygous plants (Figure 6.8).

### 6.2.7.2 Molecular and phenotypic characterization of *Arabidopsis gaza1* T-DNA insertion mutants

The *Arabidopsis gaza1* T-DNA insertion mutant seed was obtained from Dr.

Marco-Pascual

selected and

2.5) species

for a population

Systematic

100% of the 7 of them to have and the wild-type allele 15 of

heterozygous *AtRAD51/Atgad51* plant was also undertaken. There is no

phenotypic difference between the wild-type/heterozygous/homozygous

mutants in terms of the phenotypic characters.

Figure 6.7 PCR determination of the genotypes of progeny derived from a

self-fertilized heterozygous *AtRAD51/Atgad51* plant. Wild-type

and mutant alleles are indicated by red arrows. Amplification

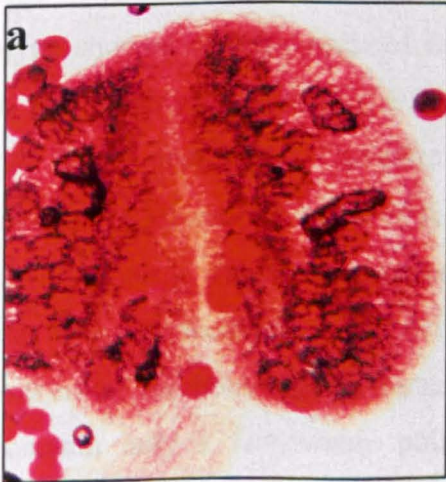
with primers for wild-type and mutant alleles identifies the

genotype of the sample analyzed; 1-25: progeny from a self-

fertilized heterozygous plant; M, 2-log ladder.

Figure 6.8 Pollen viability assay with Alexander's stain in *AtRAD51* wild-

type (a) and *Atgad51* T-DNA homozygous mutant (b).



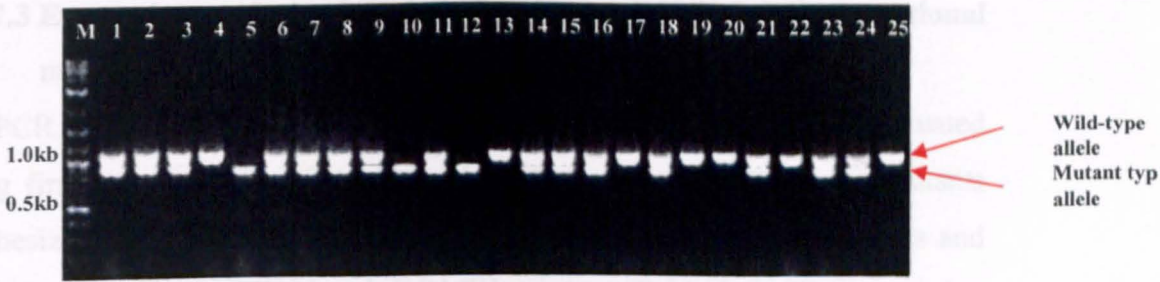
**Figure 6.8** Pollen viability assay with Alexander's stain in *AtRAD51* wild-type (a) and *Atgad51* T-DNA homozygous mutant (b).



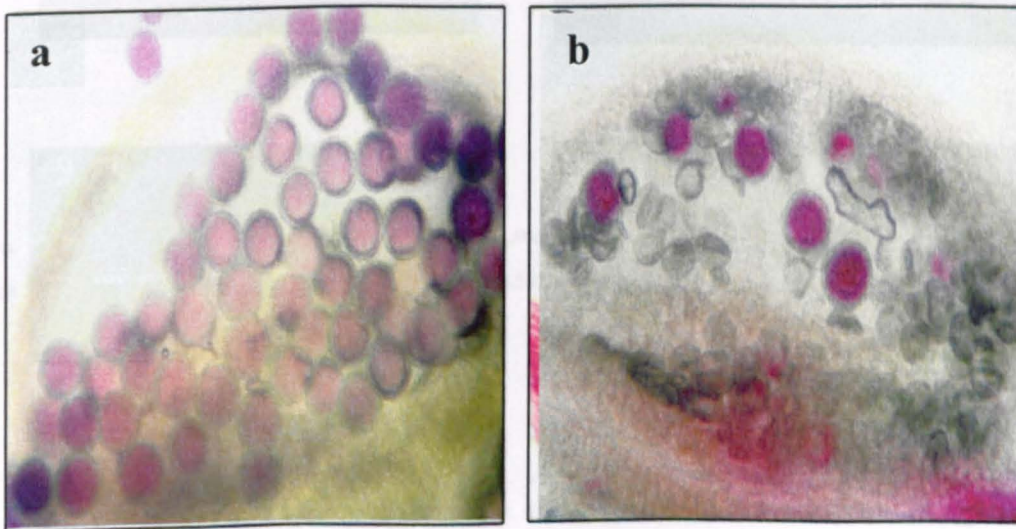
### 6.2.7.2 Molecular and phenotypic characterization of *Arabidopsis dmc1* T-DNA insertion mutants

The *Arabidopsis dmc1* T-DNA insertional mutant seed was obtained from Dr. Marie-Pascale Doutriaux. One heterozygous *AtDMC1/Atdmc1* plant was selected and systemic characterization by PCR analysis using primers (Table 2.5) specific for either the mutant or the wild-type *AtDMC1* alleles was done for a population of 25 plants derived from the self-fertilized heterozygous plant. Systematic analysis of 25 plants derived from a single self-fertilized *AtDMC1/Atdmc1* showed 7 of them to have just the wild-type alleles, 15 of them to have both alleles (indicating a heterozygote) and 3 of them to have the mutant allele to produce an approximate 1:2:1 ratios of the wild-type/heterozygous/homozygous plants ( $\chi^2=2.879$ ;  $P=0.237$ ) (Figure 6.9).

Phenotypic characterization of progeny derived from a self-fertilized heterozygous *AtDMC1/Atdmc1* plant was also undertaken. There is no phenotypic difference between the wild-type/heterozygous/homozygous mutants in terms of vegetative development. Plants with the *atdmc1/atdmc1* genotype showed strongly reduced fertility (17.5%), whereas plants with the other two genotypes *AtDMC1/AtDMC1* and *AtDMC1/Atdmc1* showed the normal fertile phenotype. The phenotypic ratio of the 25 plants gave an approximately 3:1 ratio in terms of fertile to sterility (22 fertile: 3 sterile;  $\chi^2 = 1.40$ , 1df = n.s.). The *Atdmc1/Atdmc1* plants produce smaller and curved siliques with few seeds compared to the wild-type and heterozygous plants. In the anthers of the homozygous mutant plants, shortened filament were found and contain only a few viable pollen grains, compared to wild-type and heterozygous plants, where the anthers contain normal filaments and viability of pollen grains was much higher (Figure 6.10).



**Figure 6.9** PCR determination of the genotypes of progeny derived from a single self-fertilized heterozygous *AtDMC1/Atdmc1* plant. The top band represents the wild-type allele and the bottom band represents the mutant allele. Amplification with specific primers for wild-type and mutant alleles gives the genotype of the sample analyzed; 1-25: progeny from a self-fertilized heterozygous plant; M, 2-log ladder.

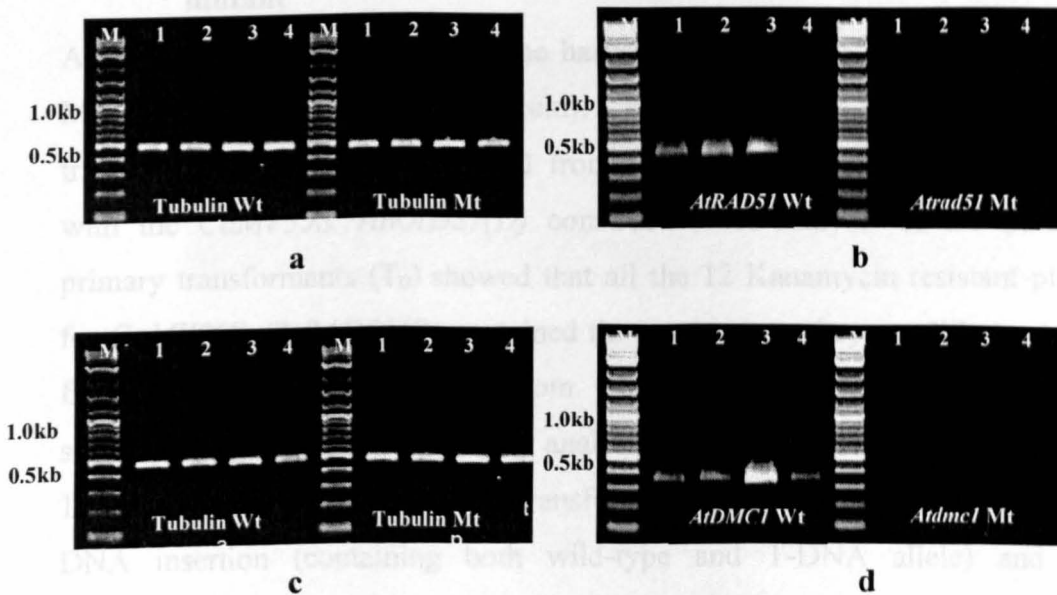


**Figure 6.10** Pollen viability staining with Alexander's solution in *AtDMC1* wild-type (a) and *Atdmc1* homozygous mutant (b).

### 6.2.7.4 Genetic transformation of *CaMV35S::TubRAD51(D)::GFP*

#### 6.2.7.3 Expression analysis of the *AtRAD51* and *AtDMC1* T-DNA insertional mutants

RT-PCR expression analysis of *AtRAD51* and *AtDMC1* genes was evaluated using first strand cDNA of Arabidopsis wild-type and homozygous mutants synthesized from the following tissues: leaf, root, floral bud and siliques and amplified with the *AtRAD51* and *AtDMC1* primers (Table 2.5). The expression of the *AtRAD51* and *AtDMC1* genes was found in leaf, root, floral bud tissues, and siliques in the wild-type genotype (Figure 6.11b and 6.11d). No expression was found in any of the tissues with the above primers in the homozygous mutant genotype, indicating the loss or lack of accumulation of the transcripts for the *AtRAD51* and *AtDMC1* gene (Figure 6.11b and 6.11d).



**Figure 6.11** RT-PCR to determine the expression of the Tubulin gene in wild-type and mutant plants of *Atrad51* and *Atdmcl* insertional lines (a and c), expression of the *AtRAD51* gene in wild-type and mutant plants in *Atrad51* insertional lines (b) and expression of *AtDMC1* gene in wild-type and mutant plants of *Atdmcl* insertional lines (d); 1-Leaf, 2-root, 3-flower and 4-silique; M, 2-log ladder.

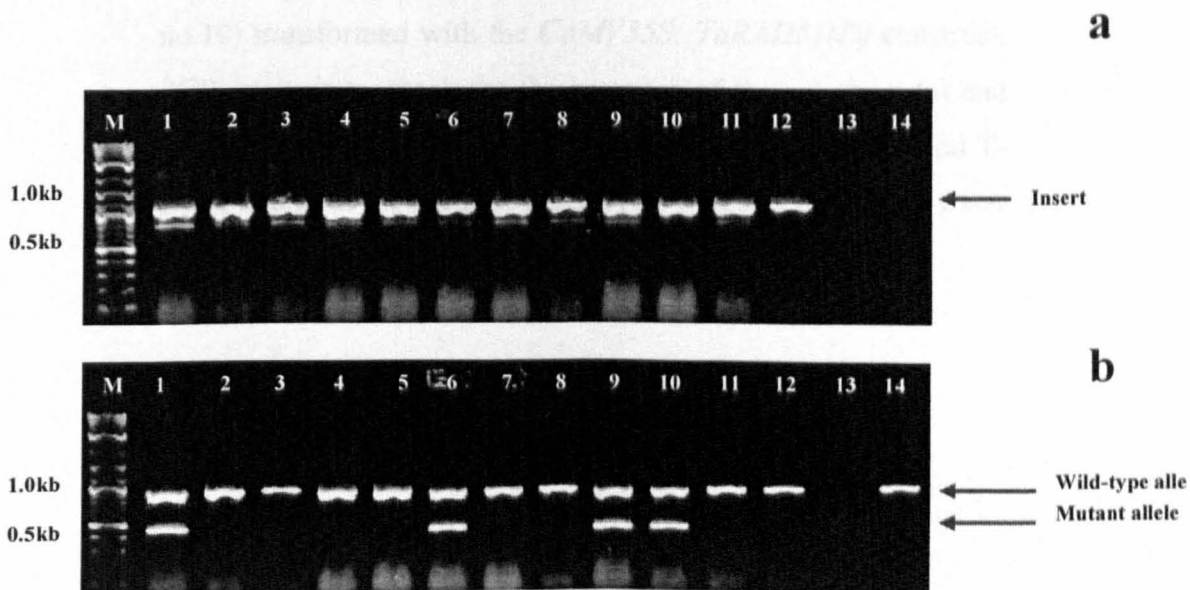
#### **6.2.7.4 Genetic transformation of *CaMV35S::TaRAD51(D)::±GFP* complementation constructs into the *Atrad51* T-DNA insertional mutant**

As homozygous *Atrad51/Atrad51* T-DNA insertional mutants of Arabidopsis are sterile, heterozygote plants were transformed with the *CaMV35S::TaRAD51(D)::±GFP* complementing constructs (Section 2.2.16.4). The heterozygous *AtRAD51/Atrad51* line was identified by using PCR primers that amplify the wild-type allele and the T-DNA allele (Section 6.2.3).

##### **6.2.7.4.1 Genetic transformation of the *CaMV35S::TaRAD51(D)* complementation construct into the *Atrad51* T-DNA insertional mutant**

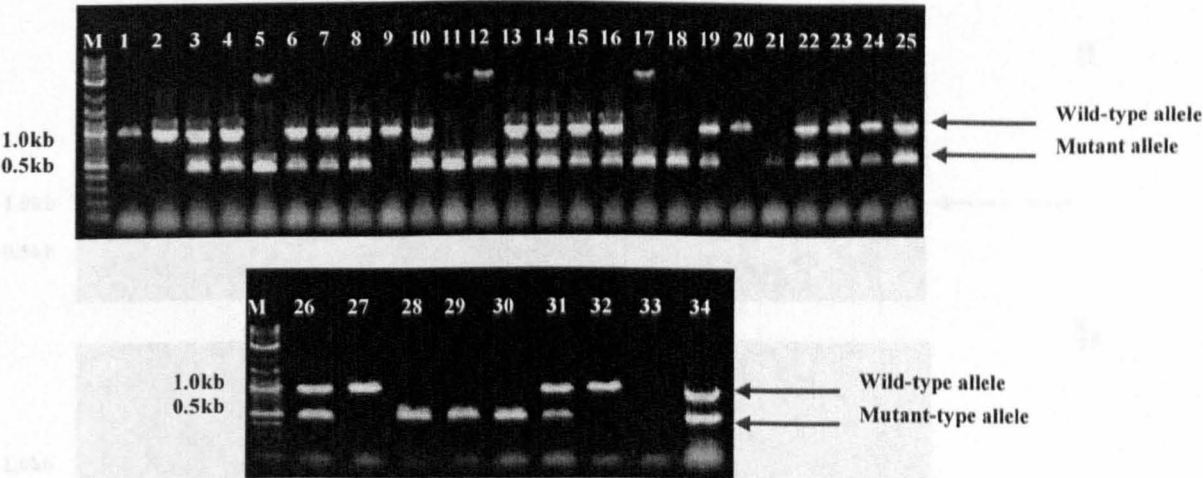
After transformation and plating the harvested seeds on selection medium ( $\frac{1}{2}$  MS media containing Kanamycin), 12 Kanamycin resistant putative transformants ( $T_0$ ) were recovered from the Arabidopsis plants transformed with the *CaMV35S::TaRAD51(D)* construct. PCR analysis of the putative primary transformants ( $T_0$ ) showed that all the 12 Kanamycin resistant plants for *CaMV35S::TaRAD51(D)* contained the construct as they amplified a 1.1kb fragment which was expected from the full length *TaRAD51(D)* cDNA sequence (Figure 6.12a). Genotypic analysis of the plants indicated that out of 12 Kanamycin resistant plants, 4 transformants were heterozygous for the T-DNA insertion (containing both wild-type and T-DNA allele) and the remaining 8 transformants contained only the wild-type allele (Figure 6.12b). PCR was repeated twice to test for consistency. As no plant with the homozygous *Atrad51/Atrad51* genotype was found, two independent heterozygous *AtRAD51/Atrad51* transformants ( $T_0$ ) (plants 1 & 10) were selected and segregated to  $T_1$  to generate homozygous *Atrad51/Atrad51* plants containing the *TaRAD51(D)* transgene to test for complementation. PCR analysis of the transformants ( $T_1$ ) from plant no.10 indicated that all 32 plants contain the integrated construct as they amplified the expected 1.1kb fragment,

probably indicative of more than one insertion of the construct into the genome at  $T_0$  for this plant. Upon genotyping with primers that amplify the wild-type allele and mutant T-DNA alleles, 18 plants were found to be heterozygous for the T-DNA, five plants were wild-type and nine plants were found be homozygous for the T-DNA insertion (Figure 6.13). For plant no.1, only 10 out of the 14  $T_1$  segregants contained the integrated construct, amplifying the expected 1.1kb fragment (Figure 6.14a). Upon genotyping with primers that amplify the wild-type allele and mutant T-DNA allele, seven plants were found to be heterozygous for the T-DNA, three plants were wild-type and three plants were found be homozygous *Atrad51/Atrad51* (PCR failed in one sample) (Figure 6.14b).

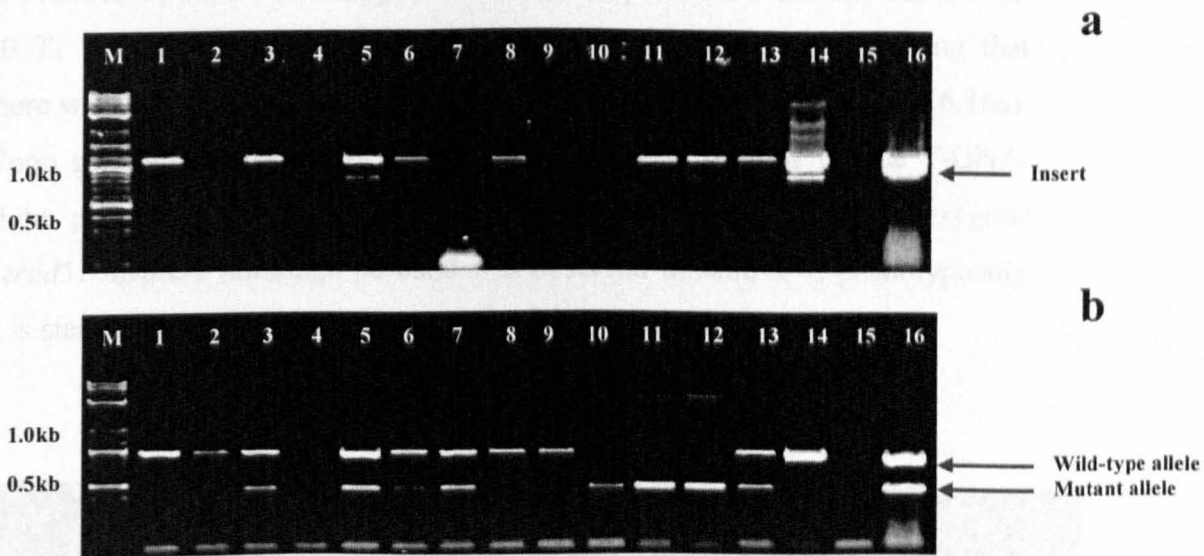


**Figure 6.12** PCR analyses of  $T_0$  plants obtained from transformation with the *CaMV35S::TaRAD51(D)* construct into *AtRAD51/Atrad51* heterozygous plants. PCR analysis was used to check for the presence of the construct (a) and PCR analysis used to genotype  $T_0$  plants for the wild-type and T-DNA alleles at the *AtRAD51* locus (b); 1-12:  $T_0$  plants, 13-water control and 14-positive control; M, 2-log ladder.





**Figure 6.13** PCR analyses of T<sub>1</sub> plants obtained from a T<sub>0</sub> transformant (plant no.10) transformed with the *CaMV35S::TaRAD51(D)* construct. PCR analysis to check for the presence of the construct (a) and PCR analysis to genotype the T<sub>1</sub> plants for the wild-type and T-DNA alleles at the *AtRAD51* gene (b); 1-32: T<sub>1</sub> plants, 33-water control and 34-positive control; M, 2-log ladder.



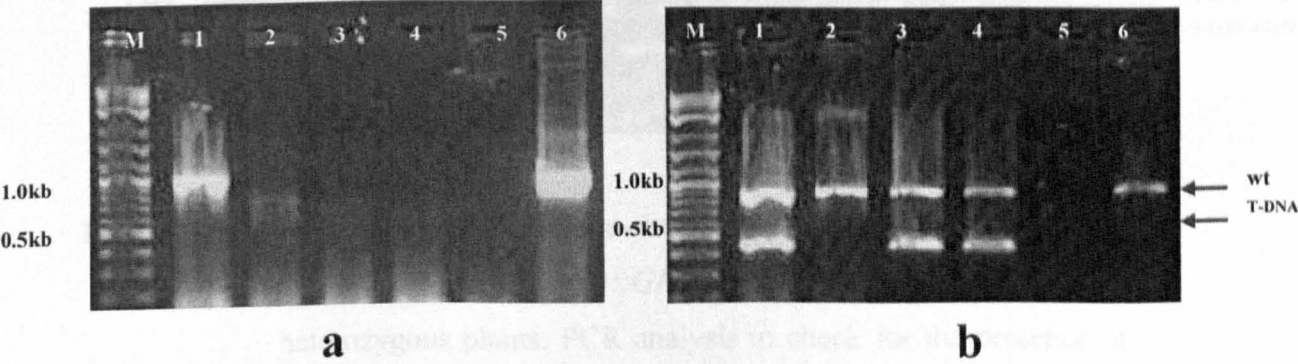
**Figure 6.14** PCR analyses of T<sub>1</sub> plants obtained from a T<sub>0</sub> transformant (plant no.1) transformed with the *CaMV35S::TaRAD51(D)* construct. PCR analysis to check for the presence of the construct (a) and PCR analysis to genotype T<sub>1</sub> plants for the wild-type and T-DNA alleles at the *AtRAD51* locus (b); 1-14: T<sub>1</sub> plants, 15-water control and 16-positive control; DNA from plant 4 failed to amplify; M, 2-log ladder

#### 6.2.7.4.2 Genetic transformation of the *CaMV35S::TaRAD51(D)::GFP* construct into the *Atrad51* T-DNA insertional mutant line

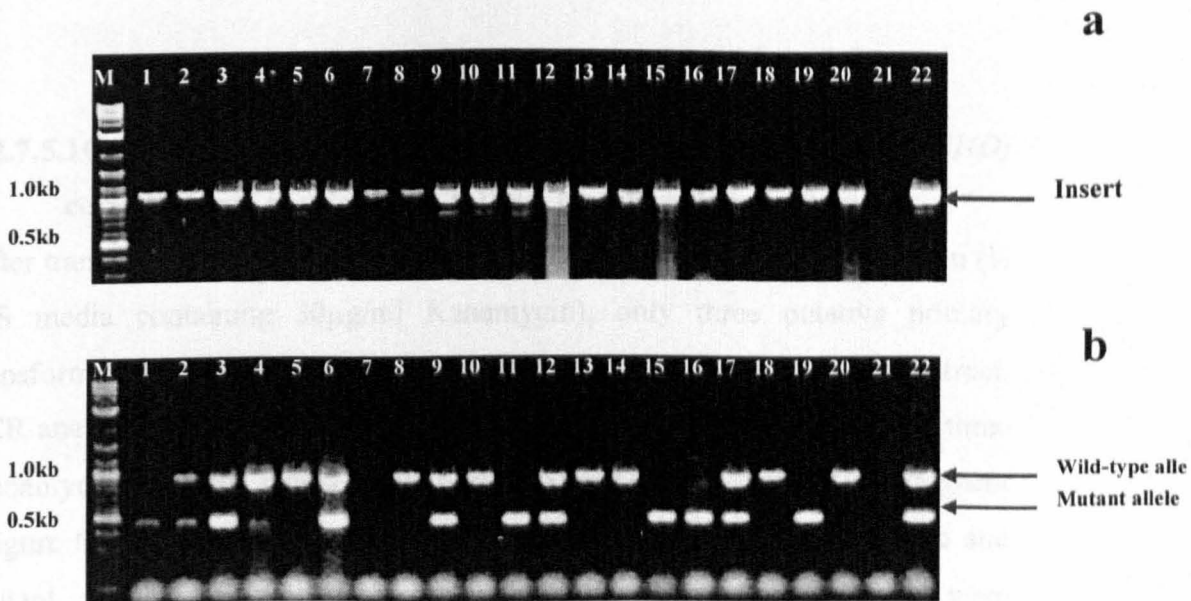
After transformation and plating the harvested seeds on selection medium ( $\frac{1}{2}$  MS media containing Kanamycin), four Kanamycin resistant primary putative transformants (T<sub>0</sub>) were recovered. PCR analysis of the primary putative transformants (T<sub>0</sub>) showed that out of four transformants only one transformant contained the expected construct fragment of 1.1kb fragment (Figure 6.15a). Genotypic analysis of the plants with *AtRAD51* wild-type and mutant T-DNA primers indicated the single primary putative transformant was heterozygous for the T-DNA insert (Figure 6.15b). The PCR was repeated twice to check for



consistency. The single transformant was selected and allowed to self-pollinate to produce T<sub>1</sub> seed. PCR analysis of the T<sub>1</sub> was performed and showed that all 20 T<sub>1</sub> transformants contained the expected 1.1kb fragment, indicating that there were probably multiple construct insertions in this T<sub>0</sub> plant (Figure 6.16a). Upon genotyping, six plants were found to be heterozygous for the T-DNA, eight plants were wild-type and six plants were found be homozygous *Atr*ad51/*Atr*ad51 (although no band was observed in sample 7, phenotypically it is sterile suggesting that it is a homozygous mutant) (Figure 6.16b).



**Figure 6.15** PCR analyses of T<sub>0</sub> plants obtained from transformation with the *CaMV35S::TaRAD51(D)::GFP* construct into *Atr*ad51 heterozygous plants. PCR analysis to check for the presence of the construct insert (a) and PCR analysis to genotype T<sub>0</sub> plants for the wild-type and T-DNA allele at the *AtRAD51* locus (b); 1-4: T<sub>0</sub> plants, 5-water control and 6-positive control; M, 2-log ladder.



**Figure 6.16** PCR analyses of T<sub>1</sub> plants obtained from transformation with the *CaMV35S::TaRAD51(D)::GFP* construct into *Atrad51* heterozygous plants. PCR analysis to check for the presence of the construct insert (a) and PCR analysis to genotype T<sub>1</sub> plants for the wild-type and T-DNA alleles at the *AtRAD51* locus (b); 1-20: T<sub>1</sub> plants, 21-water control and 22-positive control; M, 2-log ladder.

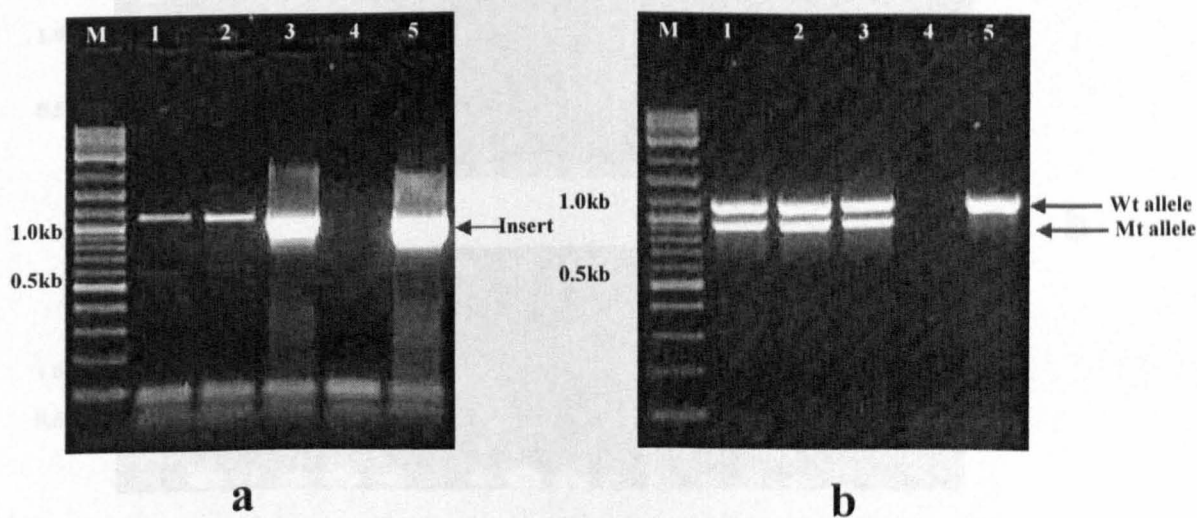
#### 6.2.7.5 Genetic transformation of *CaMV35S::TaDMC1(D)::±GFP* complementation constructs into the *AtDMC1* T-DNA insertional mutation line

As the homozygous *Atdmc1/Atdmc1* T-DNA insertional mutants are only partially fertile, genetic transformation of the *CaMV35S::TaDMC1(D)::±GFP* constructs were done into the heterozygous *AtDMC1/Atdmc1* insertion mutants. The heterozygous *AtDMC1/Atdmc1* line was identified by using PCR primers

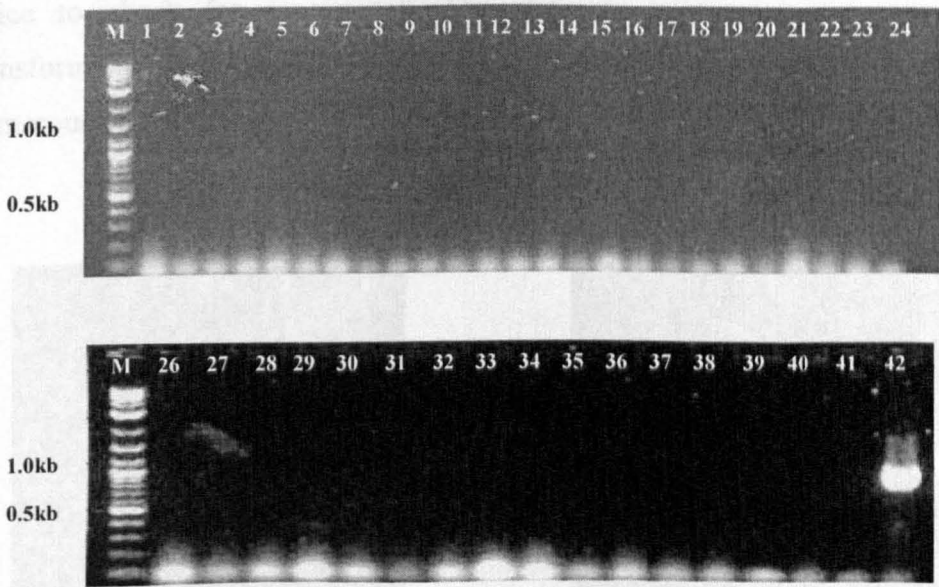
that amplify a portion on the wild-type allele and a portion on the T-DNA allele.

#### **6.2.7.5.1 Genetic transformation of *CaMV35S::TaDMC1(D)* construct into the *AtDMC1* T-DNA insertional mutant line**

After transformation and plating of the harvested seeds on selection medium ( $\frac{1}{2}$  MS media containing 30 $\mu$ g/ml Kanamycin), only three putative primary transformants ( $T_0$ ) were identified with the *CaMV35S::TaDMC1(D)* construct. PCR analysis of the three putative primary transformants showed that all three Kanamycin resistant plants contain the expected construct 1.1kb fragment (Figure 6.17a). Genotypic analysis of the plants with *AtDMC1* wild-type and mutant T-DNA primers indicated that all three transformants were heterozygous for T-DNA insert (Figure 6.17b). The PCR was repeated twice for consistency. One heterozygous *AtDMC1/Atdmc1* transformant with the *TaDMC1(D)* transgene (plant no. 3) was selected and allowed to self-pollinate to produce  $T_1$  plants. PCR analysis of the  $T_1$  indicated that all 40  $T_1$  plants did not contain the expected 1.1Kbp fragment indicative of the construct (Figure 6.18). This result might be explained by the failure to select the  $T_1$  segregants on selection media ( $\frac{1}{2}$  MS +Kan). The same transformant ( $T_0$ , plant no.3) was again plated on selection medium. No plants with Kanamycin resistance were recovered among  $T_1$  segregants on selection medium. No plants with Kanamycin resistance were found with the other two ( $T_0$ ) transformants also (Plant no. 1 & 2). The attempts to recover more primary transformants with *CaMV35S::TaDMC1(D)* construct failed despite numerous attempts. The primary reasons could be that the original transformation constructs may not have been stably integrated in the genome or that the transformants were strongly selected against *in planta* or the two T-DNA inserts apart from the original *AtDMC1* T-DNA in the Fieldman line have Kanamycin marker in them and the current construct have also the same marker and so the selection was difficult.



**Figure 6.17** PCR analyses of T<sub>0</sub> plants obtained from transformation with the *CaMV35S::TaDMC1(D)* construct into the *AtDMC1/Atdmc1* heterozygous T-DNA insertion plants. PCR analysis to check for the presence of the construct (a) and PCR analysis to genotype plants for the wild-type and T-DNA alleles at the *AtDMC1* locus (b); 1-3: T<sub>0</sub> plants, 4-water control and 5-positive control; M, 2-log ladder.



b

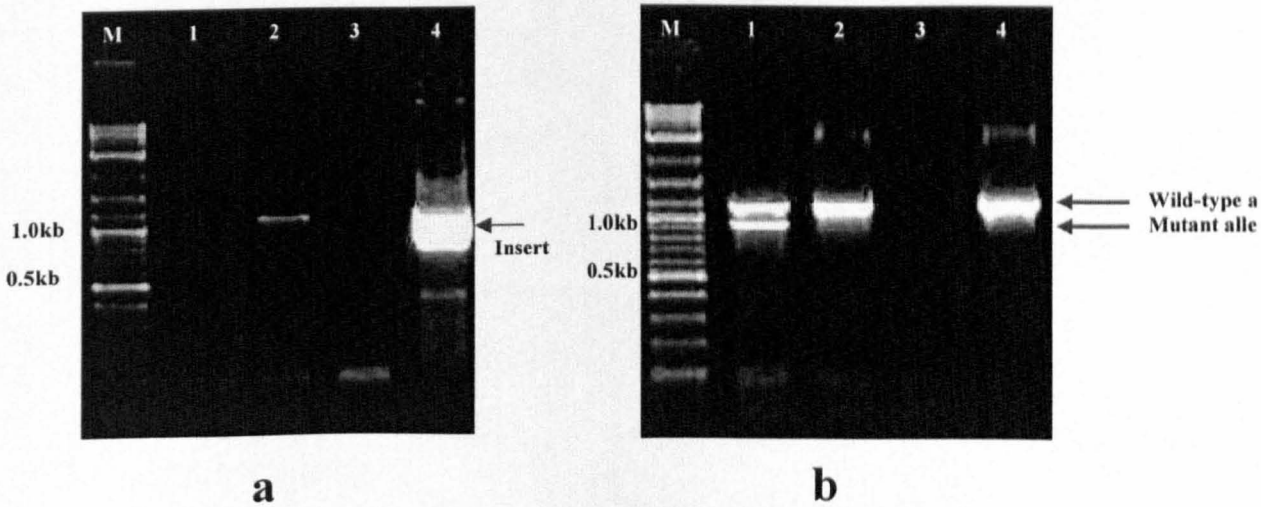
**Figure 6.18** PCR analyses of  $T_1$  plants obtained from transformation with the *CaMV35S::TaDMC1(D)* construct into *AtDMC1* T-DNA heterozygous plants. PCR analysis to check for the presence of the insert (a) and PCR analysis to genotype  $T_1$  plants for the wild-type and T-DNA alleles at the *AtDMC1* locus (b); 1-40:  $T_1$  plants, 41-water control and 42-positive control; M, 2-log ladder.

#### 6.2.7.5.2 Genetic transformation of the *CaMV35S::TaDMC1(D)::GFP* construct into the *AtDMC1* T-DNA insertional mutant line

After transformation and plating of the harvested seeds on selection medium ( $\frac{1}{2}$  MS media containing 30  $\mu\text{g/ml}$  Kanamycin), only two putative transformants ( $T_0$ ) were recovered with the *CaMV35S::TaDMC1(D)::GFP* construct. PCR analysis of the two putative primary transformants ( $T_0$ ) showed that only one Kanamycin resistant plant contained the correct construct, amplifying the expected 1.1kb fragment (Figure 6.19a). Genotypic analysis of the plants with *AtDMC1* wild-type and T-DNA mutant primers indicated that the single ( $T_0$ ) transformant has wild-type alleles only (Figure 6.19b). The PCR was repeated



twice to check for consistency. The attempts to recover more primary transformants with the *CaMV35S::TaDMC1(D)::GFP* construct failed despite numerous attempts.

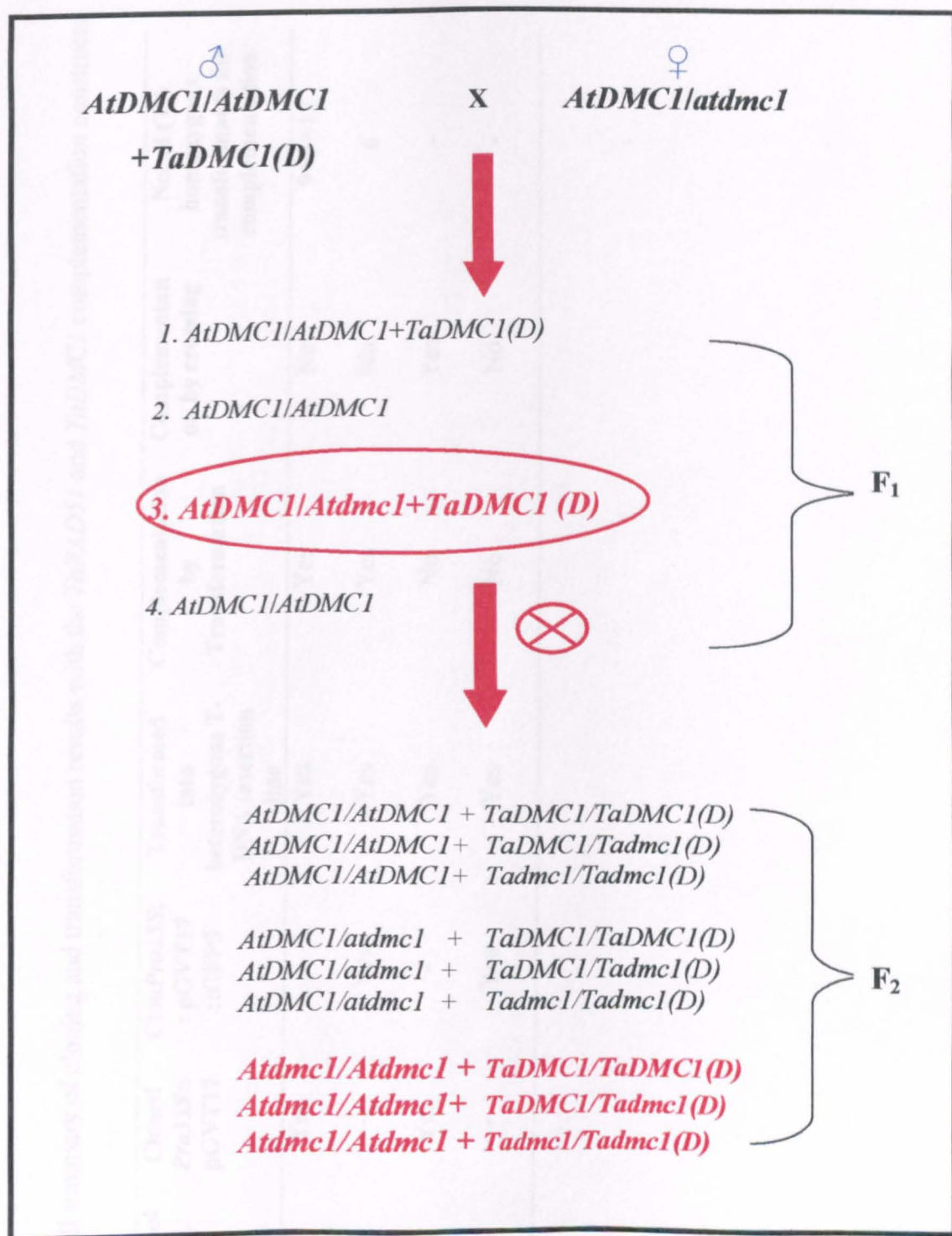


**Figure 6.19** PCR analyses of  $T_0$  plants obtained from transformation with the *CaMV35S::TaDMC1(D)::GFP* construct into *AtDMC1* T-DNA insertion heterozygous plants. PCR analysis to check for the presence of the construct insert (a) and PCR analysis to genotype plants for the wild-type and T-DNA alleles at the *AtDMC1* locus (b); 1-2:  $T_0$  plants, 3-water control and 4-positive control; M, 2-log ladder.

In the absence of any primary transformant ( $T_0$ ) with either a heterozygous or homozygous *Atdmc1* background, the primary *CaMV35S::TaDMC1(D)::GFP* transformant in an Arabidopsis wild-type background was crossed to a heterozygous *AtDMC1/Atdmc1* plant (Figure 6.20). The  $F_1$  progeny that are heterozygous for *AtDMC1/Atdmc1* expressing the transgene were produced but there was insufficient time to analyze the  $F_2$  progeny in this work. The presence of the GFP fusion is also a potential cause for concern as there is currently no

plant containing the construct for over-expressing *TaDMC1(D)* without this fusion.





**Figure 6.20** Diagram depicting the crossing strategy employed to transfer the *TaDMC1(D)* transgene from the wild-type background to a *Atdmc1/Atdmc1* homozygous mutant background in Arabidopsis.

**Table 6.1a** Overall summary of cloning and transformation results with the *TaRAD51* and *TaDMC1* complementation constructs

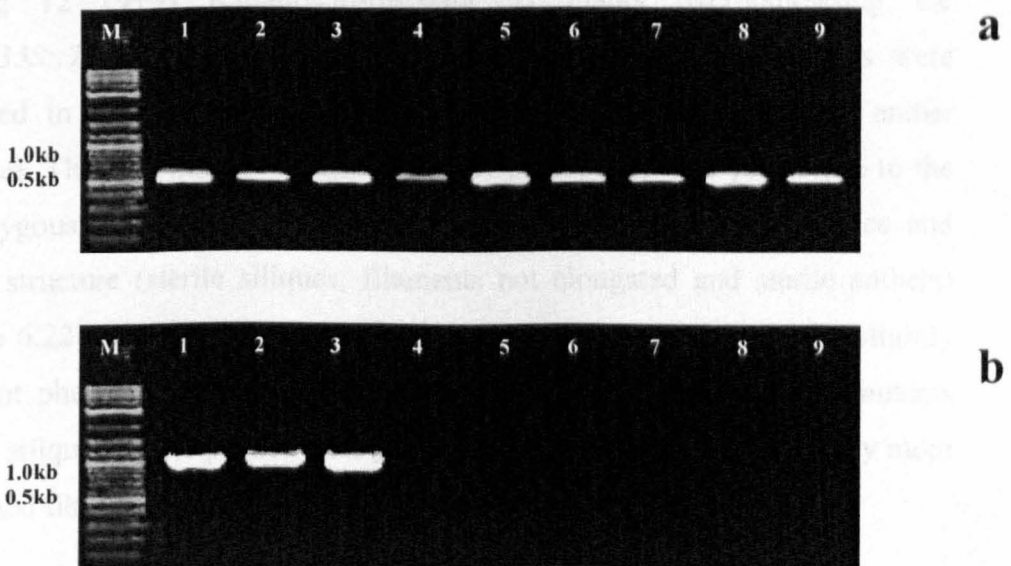
Gene	Homoeologue	Cloned <i>Pro35S</i> :: pGVT17	Clon <i>Pro35S</i> : : pGVT17 ::GFP5	Transformed into heterozygous T- DNA insertion line	Complementation by Transformation	Complementati on by crossing	No. of (T <sub>1</sub> ) homozygous transformants for complementation
<i>TaRAD51</i>	(D)	Yes	-	Yes	Yes	No	9+3=12
<i>TaRAD51</i>	(D)	-	Yes	Yes	Yes	No	6
<i>TaDMC1</i>	(D)	Yes	-	Yes	No	Yes	-
<i>TaDMC1</i>	(D)	-	Yes	Yes	No	No	-

**Table 6.1b** Overall summary of cloning and transformation results with the *TaRAD51* and *TaDMC1* Complementation constructs to date

Gene	Construct	No. of T <sub>0</sub> recovered on selection medium	No. of T <sub>0</sub> used for complementa tion studies	No. of T <sub>1</sub> generated from each T <sub>0</sub>	T <sub>1</sub> segregation for construct- insert by PCR
<i>TaRAD51(D)</i>	<i>CaMV35S</i>	12	2	40 and 14	40/40 and 9/14
<i>TaRAD51(D)</i>	<i>CaMV35S::GFP</i>	1	1	20	20/20
<i>TaDMC1(D)</i>	<i>CaMV35S</i>	3	1	40	0/40
<i>TaDMC1(D)</i>	<i>CaMV35S::GFP</i>	1	1	-	-

#### 6.2.7.6 Confirmation of *TaRAD51(D)* expression in *Atrad51* homozygous transformed plants

Expression analysis by RT-PCR showed that the *TaRAD51(D)* gene produced a transcript in leaf, flower and siliques of one of the plants (Plant no.12) transformed with the *CaMV35S::TaRAD51(D)::GFP* construct but no bands were found in non-transformed, wild-type or homozygous mutant plants, as expected (Figure 6.21b). This also confirms that the transcripts from the *TaRAD51(D)* can be distinguished clearly from the endogenous Arabidopsis version. This results also confirmed that all the transformant plants express the *TaRAD51(D)* gene when the relevant construct was present (checked by genomic PCR for construct insert) and hence these plants could be used to study any potential functional complementation effects.



**Figure 6.21** RT-PCR expression analysis of transformant (plant no.12) of *CaMV35S::TaRAD51(D)::GFP* construct with Tubulin primers (a) and with the *TaRAD51* gene primers (to amplify ORF) (b); 1-leaf, 2-silique and 3-flower from the transformant plants; 4-leaf, 5-flower and 6-silique from the *AtRAD51* wild-type plants; 7-leaf, 8-silique and 9-flower from *AtRAD51* T-DNA mutant plants; M, 2-log ladder.

### **6.2.7.7 Complementation of the *Atrad51* insertional mutation line with the over-expressing *CaMV35S::TaRAD51(D)::±GFP* construct**

#### **6.2.7.7.1 Complementation with the *CaMV35S::TaRAD51(D)* construct**

From 46 (32+14) T<sub>1</sub> transformants derived from the T<sub>0</sub> plant by segregation only 41 were found to carry the over-expressing constructs and the remaining five did not have it. Among these 46 transformants, 34 were found to contain at least one wild type allele (*AtRAD51*), while the other 12 plants did not contain any wild type alleles. To examine the ability of the *TaRAD51* gene product to functionally complement the homozygous mutants caused by disruption of *AtRAD51* gene, these 12 homozygous T-DNA mutant lines carrying the over-expression *CaMV35S::TaRAD51(D)* construct were phenotypically analyzed.

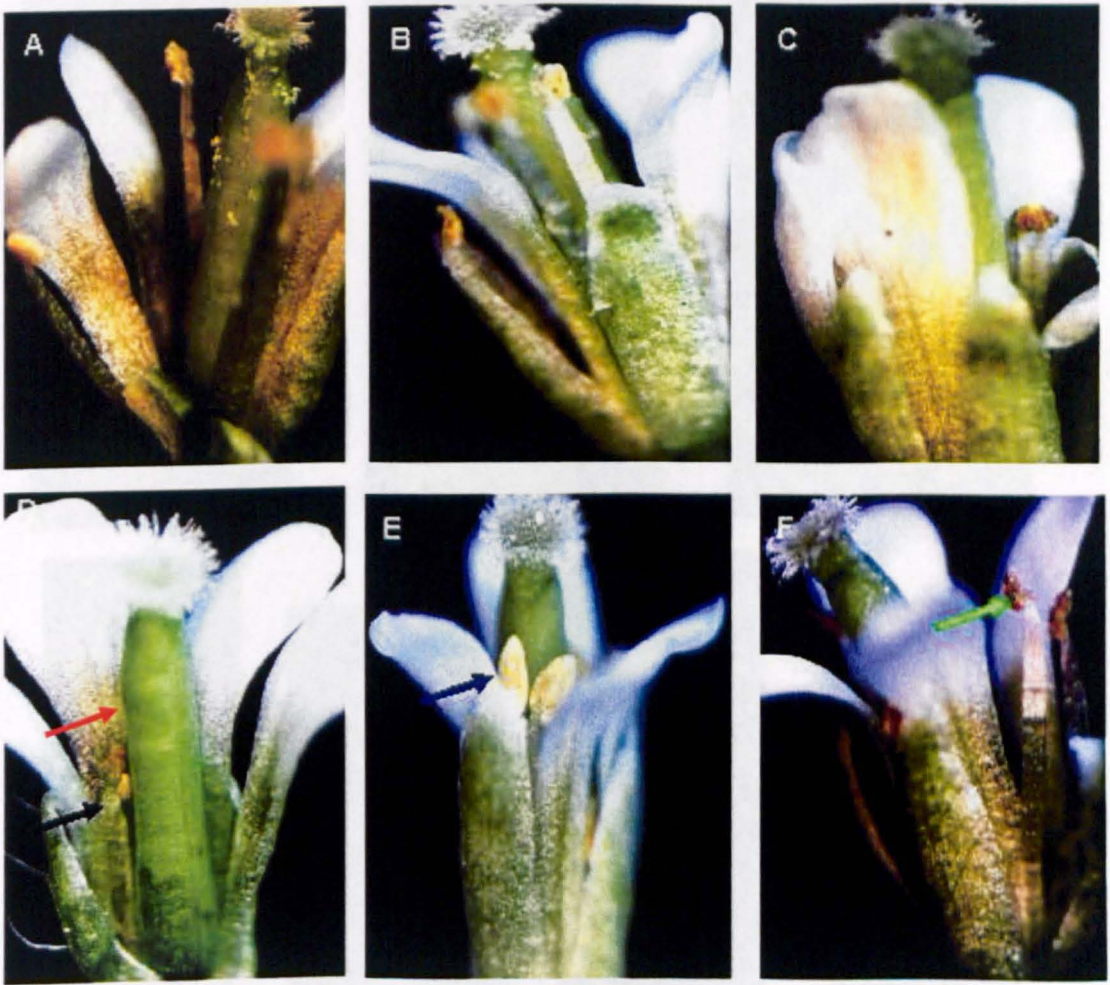
Among 12 (9+3) Kanamycin-resistant T<sub>1</sub> plants over-expressing the *CaMV35S::TaRAD51(D)* construct two different kinds of phenotypes were observed in the plants, particularly in the inflorescence and in the anther structure. The first kind of transformants displayed a similar phenotype to the homozygous *Atrad51/Atrad51* mutants with regard to the inflorescence and anther structure (sterile siliques, filaments not elongated and sterile anthers) (Figure 6.22 and 6.23). The second kind of transformants displayed a slightly different phenotype compared to the homozygous *Atrad51/Atrad51* mutants (fertile siliques on the plant, viable pollen grains in the anther and slightly more elongated filaments) (Figure 6.22 and 6.23)

#### **6.2.7.7.2 Complementation with the *CaMV35S::TaRAD51(D)::GFP* construct**

All 20 T<sub>1</sub> transformants derived from the single T<sub>0</sub> transformant by segregation were found to carry the over-expressing constructs. Among these 20 transformants, 14 were found to contain at least one wild type allele (*AtRAD51*), while the other six plants did not contain any wild type alleles. To examine the ability of the *TaRAD51* gene product to functionally complement

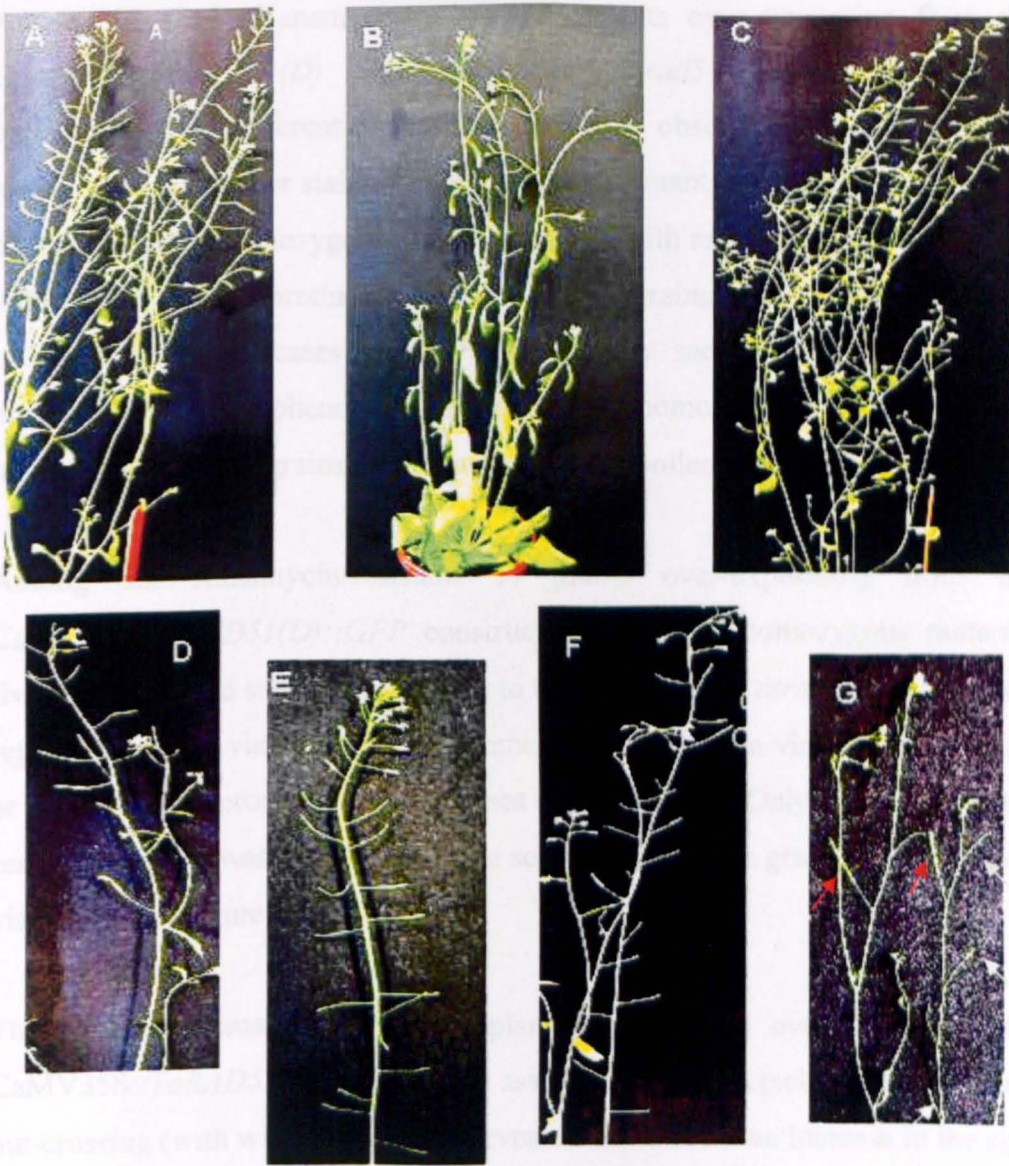
the homozygous mutants caused by disruption of *AtRAD51* gene, these six homozygous T-DNA mutant lines carrying the over-expression construct were analyzed phenotypically. All the lines displayed the same phenotype as the homozygous *Atrad51/Atrad51* mutant with regard to inflorescence and anther structure (for example sterile siliques, filaments not elongated and sterile anthers) (Figure 6.22 and 6.23).





**Figure 6.22** Flower morphology of the wild-type (A and B), the *Atrad51* homozygous mutant (C) and the *Atrad51* homozygous transformant over-expressing *CaMV35S::TaRAD51(D)* (D, E and F). Red and black arrows indicate fertile and sterile anthers (mixed-type anthers) in the same flower, blue arrow indicates sterile anthers (mutant-type anthers) and green arrow indicates fertile anther (wild-type anthers).





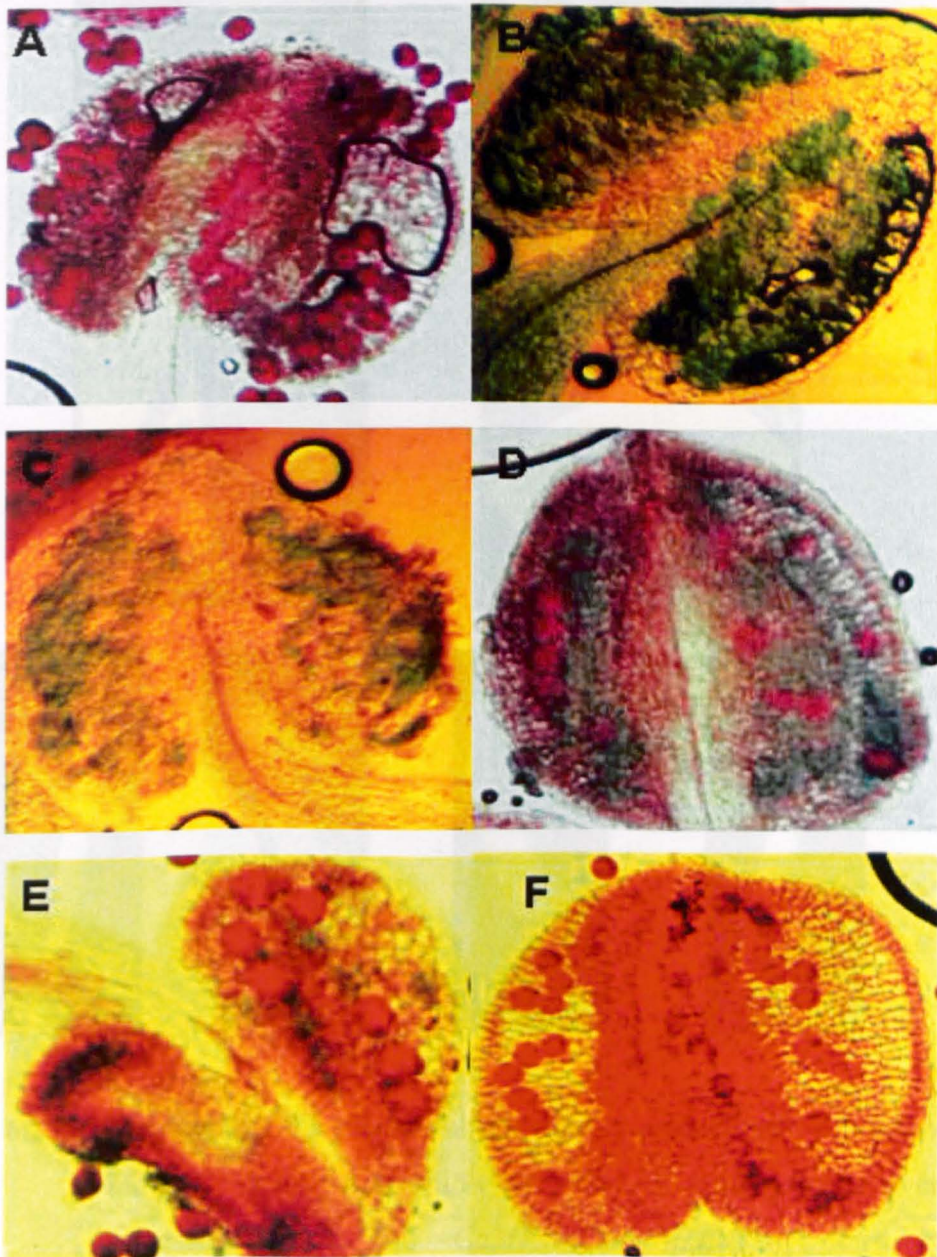
**Figure 6.23** The plant habit and inflorescence structure of the *AtRAD51* wild-type plant (A and D), *Atrad51* homozygous mutant plant (B and E) and *Atdmc1* homozygous mutants over-expressing *CaMV35S::TaRAD51(D)* (C and F; G and H). Plant (G) showed a mixed inflorescence (fertile siliques are represented by red arrows and sterile siliques by white by white arrows) with the *CaMV35S::TaRAD51(D)* construct.

Among 12 (9+3) Kanamycin-resistant T<sub>1</sub> plants over-expressing from the *CaMV35S::TaRAD51(D)* construct in an *Atrad51* homozygous mutant background, two different phenotypes were also observed on pollen viability tested using Alexander staining. The first transformant class displayed a similar phenotype to the homozygous *atrad51* mutants with regards to pollen viability, with all the anthers produce non-viable pollen grains or no pollen at all was produced in some cases (Figure 6.24C). The second transformant class displayed a different phenotype compared to the homozygous *atrad51* mutants, having viable pollen grains as well as non-viable pollen (Figure 6.24D and E).

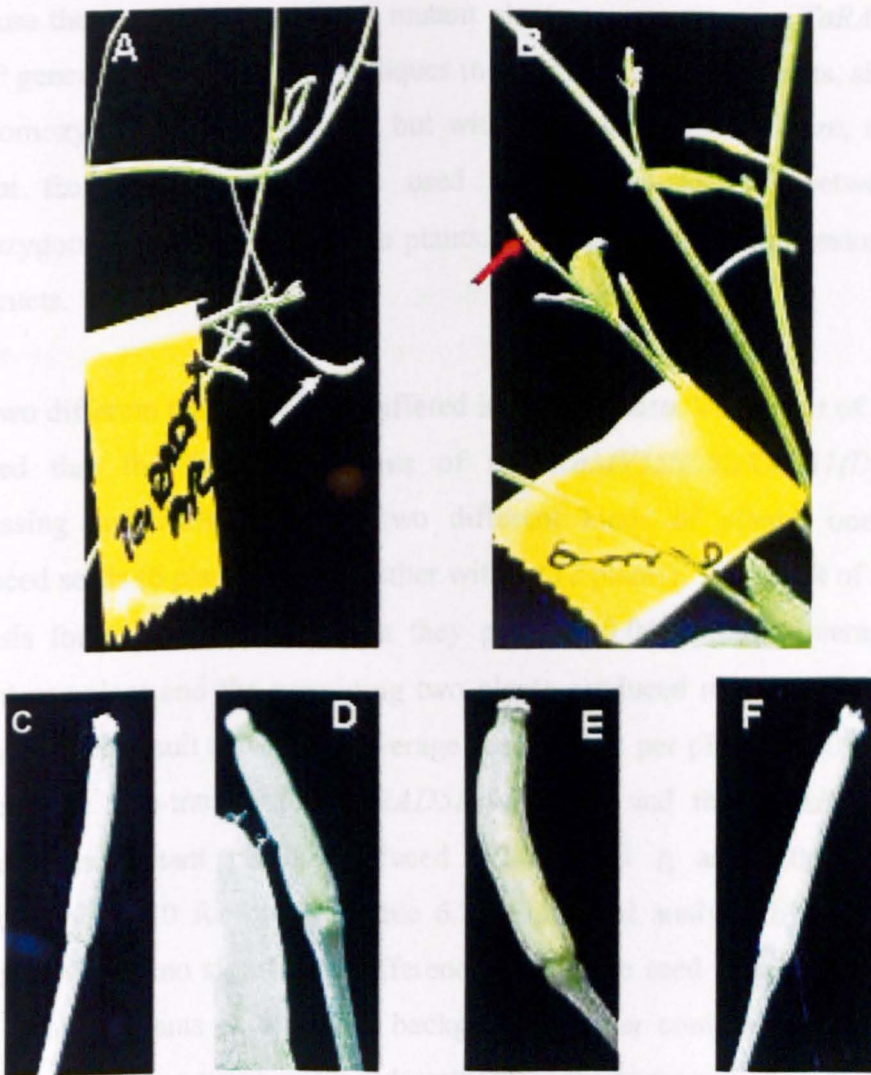
Among six Kanamycin-resistant T<sub>1</sub> plants over-expressing from the *CaMV35S::TaRAD51(D)::GFP* constructs in *Atrad51* homozygous mutants, five plants showed similar phenotypes to the homozygous *atrad51* mutants with regards to pollen viability, with all anthers producing non-viable pollen grains or no pollen was produced in some cases (Figure 6.24C). Only one Kanamycin-resistant T<sub>1</sub> plant was found to produce some viable pollen grain as well as non-viable pollen (Figure 6.24F)

The results of crossing of the T<sub>1</sub> plant (plant no.12) over-expressing the *CaMV35S::TaRAD51(D)* construct by assisted pollination (self-pollination) and out-crossing (with wild-type pollen) revealed that there is an increase in the size of the silique in assisted pollination compared to normal silique (Figure 6.25). The out-crossing of the transformant with wild-type pollen revealed an increase in silique size compared to self-assisted and normal silique (Figure 6.25).





**Figure 6.24** Pollen viability assay with Alexander's stain in wild-type (A), the *Atrad51* homozygous mutant (B) and the *Atrad51* homozygous mutants over-expressing *CaMV35S::TaRAD51* (C, D, E and F).



**Figure 6.25** Phenotype of the siliques observed after assisted self-pollination and out-crossing with wild-type pollen of transformants over-expressing from the *CaMV35S::TaRAD51(D)* construct. Crossing of transformant plant with wild type pollen (A) and assisted self-pollination of the transformant (B). Microscopic images of non-selfed silique (C), siliques obtained by crossing the transformant with wild-type pollen (D and E) and silique obtained by assisted self-pollination of a transformant (F).

Because the *atrad51* homozygous mutant plants over-expressing *TaRAD51(D)*  $\pm$ GFP generally produced small siliques in most of the transformants, similar to the homozygous *Atrad51* mutant, but with some variability of size, the seed weight from transformants was used to compare fertility between the homozygous mutant and wild-type plants, with and without expression of the constructs.

The two different types of plants differed in seed set also. The result of seed set showed that the  $T_1$  transformants of the *CaMV35S::TaRAD51(D)* over-expressing constructs produced two different kinds of plants: one which produced seeds (6 plants) and the other without (2 plants). The result of seed set analysis for 6 plants showed that they produce  $0.046 \pm 0.030$ g average seed weight per plant and the remaining two plants produced no seed (Table 6.2). The combined result showed an average seed weight per plant of  $0.035 \pm 0.03$ g, whereas the non-transformed *AtRAD51* wild-type and the *Atrad51/Atrad51* homozygous mutant plants produced  $0.244 \pm 0.051$  g and 0.0g per plant respectively ( $n=10$  for both), (Table 6.2). Statistical analysis by the Duncan method revealed no significant difference in average seed weight per plant of the  $T_1$  transformants in a mutant background (either complementing type or non-complementing type) and the *Atrad51/Atrad51* homozygous mutant. There was a significant difference between these and the wild-type plants (Table 6.3 and Figure 6.26). Whereas seed weight in the homozygous mutant background, not expressing the *TaRAD51(D)* gene was reduced to 0% of the wild-type plant seed weight, the seed weight in the homozygous mutant background over-expressing the *TaRAD51(D)* gene was reduced to 18% and 0% respectively for two different types of plants, with a combined average of 14% of the wild-type plant seed weight. This suggests that the over-expression of the *TaRAD51(D)* gene using the above construct is leading to some limited complementation of the mutant line.

The result for the  $T_1$  transformants of the *CaMV35S::TaRAD51(D)::GFP* over-expressing construct in the *Atrad51/Atrad51* homozygous background showed

0.005±0.001g seed weight per plant (n=6), while the non-transformed *AtRAD51* wild-type and the *Atrad51/Atrad51* homozygous mutant plants produced 0.244±0.051 g and 0.0g per plant (n=10 for both), respectively (Table 6.4). Statistical analysis by the Duncan method revealed no significant difference in the average seed weight per plant of T<sub>1</sub> transformants in a mutant background compared with the *Atrad51/Atrad51* homozygous mutants plants not expressing the construct (Table 6.5 and Figure 6.27). Both were significantly different from the wild type seed weight though. Compared to the wild-type seed weight, the mutant background over-expressing the *CaMV35S::TaRAD51(D)::GFP* produced only 2% complementation.

**Table 6.2**      Seed weight per plant in T<sub>1</sub> segregants of *Atrad51* homozygous mutants over-expressing from the *CaMV35S::TaRAD51(D)* construct compared with the *Atrad51* homozygous mutant and the *AtRAD51* wild-type plants. Transformants expressing a ‘sterile’ phenotype and transformant expressing a ‘mixed’ phenotype have been indicated.

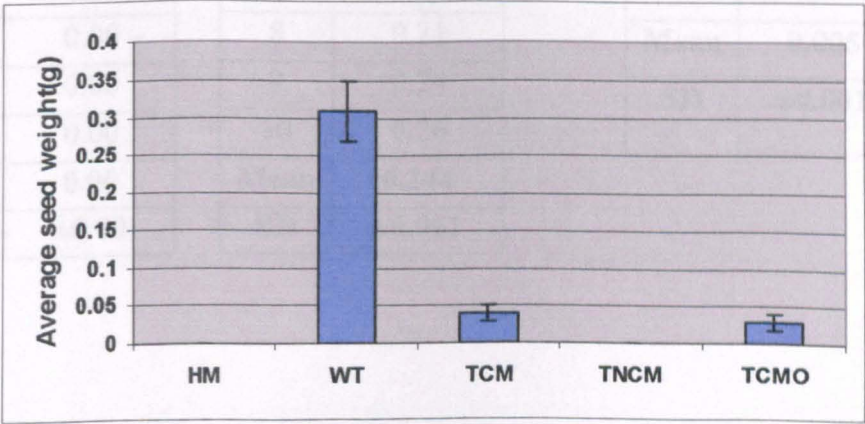
<i>Atrad51</i> homozygous mutants		<i>AtRAD51</i> wild-type plants		<i>CaMV35S::TaRAD51(D)</i> T <sub>1</sub> transformants	
Plant no.	Seed weight (g)	Plant no.	Seed weight (g)	Plant no.	Seed weight (g)
1	0.00	1	0.13	2	0.02
2	0.00	2	0.26	3	0.04
3	0.00	3	0.32	5	0.04
4	0.00	4	0.29	6	0.10
5	0.00	5	0.25	7	0.02
6	0.00	6	0.26	8	0.06
7	0.00	7	0.22	Mean	0.046
8	0.00	8	0.21	SD	±0.030
9	0.00	9	0.24	1	0.00
10	0.00	10	0.26	4	0.00
Mean	0.00	Mean	0.244	Mean	0.00
SD	±0.00	SD	±0.051	SD	0.00
				Mean of both means	0.03
				SD	±0.030



**Table 6.3** Average seed weight per plant in the T<sub>1</sub> segregants of *Atrad51* homozygous mutants over-expressing *CaMV35S::TaRAD51(D)* compared with the *Atrad51* homozygous mutant and *AtRAD51* wild-type plants.

Line name	Average seed weight (g)
<i>Atrad51</i> homozygous mutant	00.0±0.00 <sup>a</sup>
<i>AtRAD51</i> wild-type	0.31±0.04 <sup>b</sup>
<i>Atrad51</i> + <i>TaRAD51(D)</i> (Complementing type)	0.04±0.01 <sup>a</sup>
<i>Atrad51</i> + <i>TaRAD51(D)</i> (non-complementing type)	00.0±0.00 <sup>a</sup>
<i>Atrad51</i> + <i>TaRAD51(D)</i> - overall	0.03±0.01 <sup>a</sup>

Values are means of eight replicates per treatment +s.e. Values for the same parameter labeled with the same letter are not significantly different (P≤0.05)



**Figure 6.26** Bar graph showing the average seed weight per plant in T<sub>1</sub> segregants of *Atrad51* homozygous mutants over-expressing from the *CaMV35S::TaRAD51(D)* construct compared with the *Atrad51* homozygous mutant and *AtRAD51* wild-type plants; WT: *AtRAD51* wild-type plants; HM: *Atrad51* homozygous mutant; TCM: *CaMV35S::TaRAD51(D)* complementing plants in a *Atrad51/Atrad51* background; TNCM: *CaMV35S::TaRAD51(D)* non-complementing plants in a *Atrad51/Atrad51* background; TCMO: *CaMV35S::TaRAD51(D)* plants in a *Atrad51/Atrad51* background

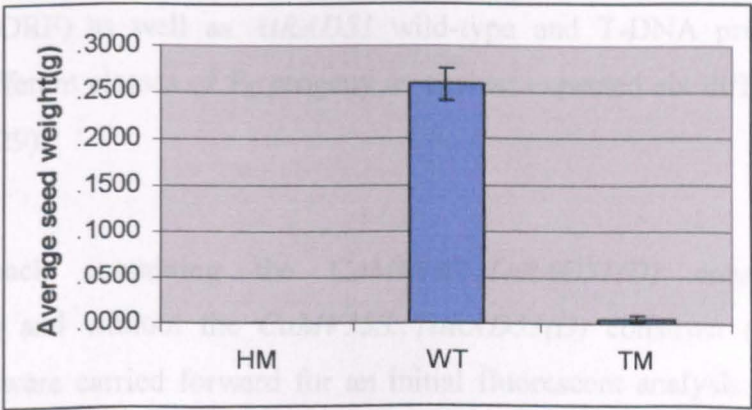
**Table 6.4**      Seed weight per plant in T<sub>1</sub> segregants of *Atrad51* homozygous mutants over-expressing *CaMV35S::TaRAD51(D)::GFP* compared with the *Atrad51* homozygous mutant and *AtRAD51* wild-type plants.

<i>Atrad51</i> homozygous mutants		<i>AtRAD51</i> wild-type plants		<i>CaMV35S::TaRAD51(D) ::GFP</i> T <sub>1</sub> transformants	
Plant no.	Seed weight (g)	Plant no.	Seed weight (g)	Plant no.	Seed weight (g)
1	0.00	1	0.13	1	0.00
2	0.00	2	0.26	2	0.00
3	0.00	3	0.32	3	0.00
4	0.00	4	0.29	4	0.00
5	0.00	5	0.25	5	0.00
6	0.00	6	0.26	6	0.02
7	0.00	7	0.22		
8	0.00	8	0.21	<b>Mean</b>	<b>0.005</b>
9	0.00	9	0.24		
10	0.00	10	0.26	<b>SD</b>	<b>±0.001</b>
<b>Mean</b>	<b>0.00</b>	<b>Mean</b>	<b>0.244</b>		
<b>SD</b>	<b>±0.00</b>	<b>SD</b>	<b>±0.051</b>		

**Table 6.5** Average seed weight per plant in T<sub>1</sub> segregants of *Atrad51* homozygous mutants over-expressing *CaMV35S::TaRAD51(D)* in a homozygous mutant background compared with the *Atrad51* homozygous mutant and *AtRAD51* wild-type plants.

Line name	Average seed weight (g)
<i>Atrad51</i> homozygous mutant	0.00±0.00 <sup>a</sup>
<i>AtRAD51</i> wild-type	0.25±0.01 <sup>b</sup>
<i>Atrad51</i> + <i>TaRAD51(D)::GFP</i>	0.00±0.00 <sup>a</sup>

Values are means of six replicates per treatment +s.e. Values for the same parameter labeled with the same letter are not significantly different (P≤0.05)



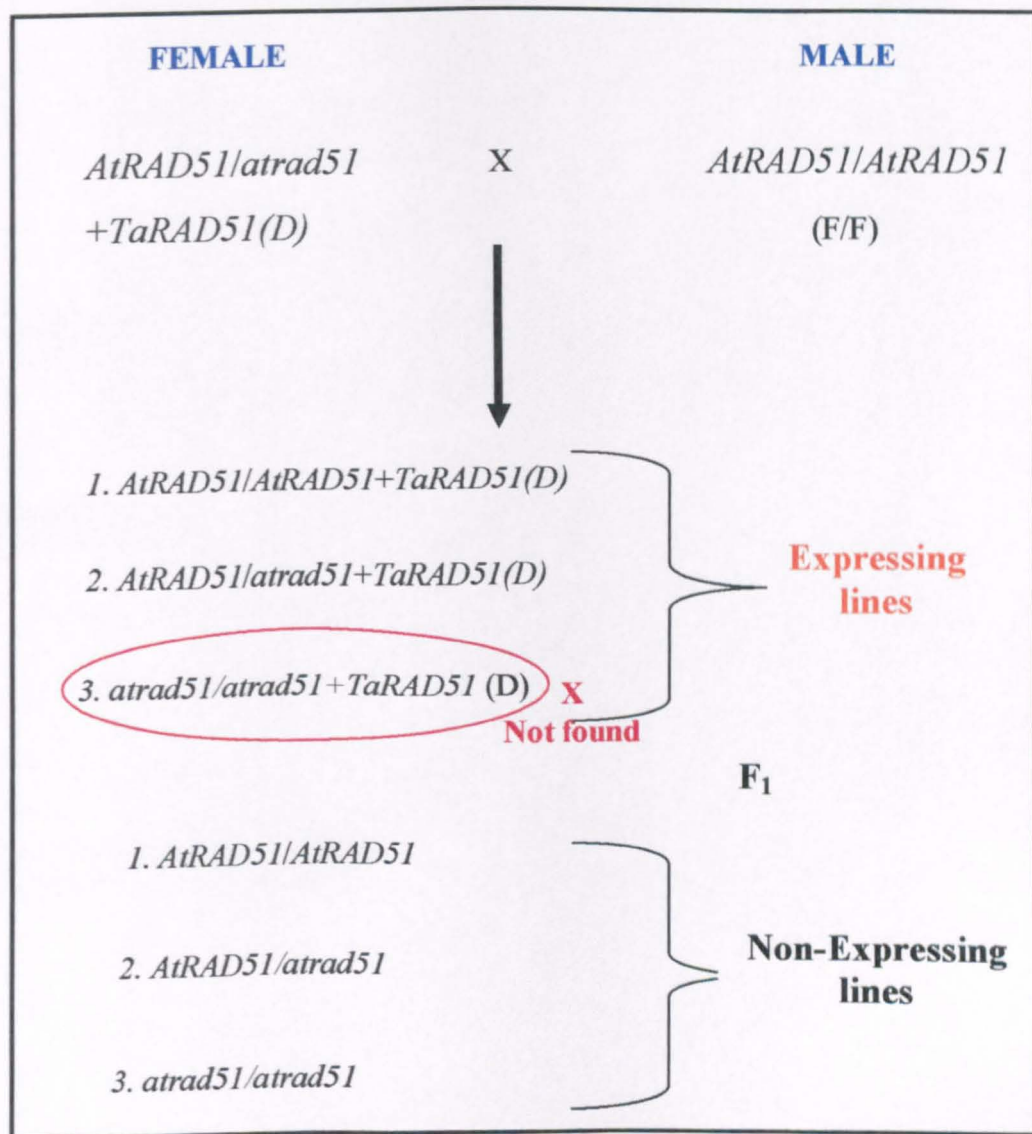
**Figure 6.27** Bar graph showing the average seed weight per plant in T<sub>1</sub> segregants of the *Atrad51* homozygous mutants over-expressing *CaMV35S::TaRAD51(D)::GFP* compared with the *Atrad51* homozygous mutant and *AtRAD51* wild-type plants; HM: *Atrad51* homozygous mutant; WT: *AtRAD51* wild-type plants; TM: *CaMV35S::TaRAD51(D)::GFP* plants in a *Atrad51/Atrad51* background.

### 6.2.8 Effect of over-expression of the *TaRAD51* gene on genetic distances and recombination frequencies

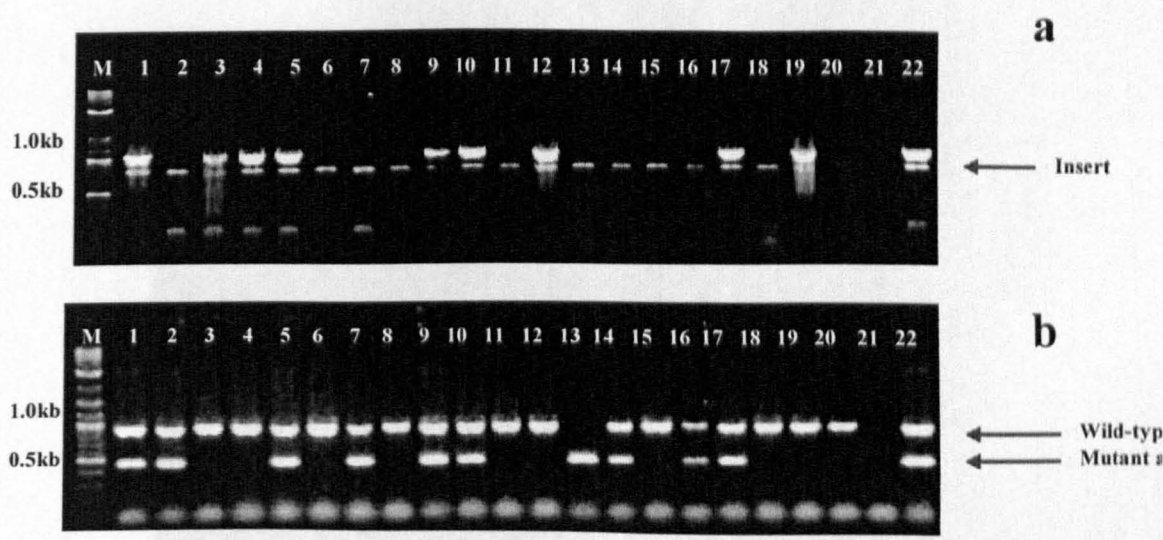
To determine genetic distances that can be visually assayed for recombination frequencies by over-expression of the *TaRAD51* transgene in Arabidopsis, heterozygous *AtRAD51/Attrad51* plants over-expressing the *CaMV35S::TaRAD51(D)* construct (used as the female plant) were crossed to Arabidopsis fluorescent tetrad line I<sub>3</sub>a (used as male) carrying differently colored markers (Red, cyan and Yellow) on the same chromosome to produce F<sub>1</sub> progeny carrying both markers in trans and creating a genetic interval bounded by two visible markers (Figure 6.28). COs were observed in the pollen grains and not pollen tetrads produced by these F<sub>1</sub> plants by visually scoring the segregation of the two fluorescent proteins (Red and Yellow) (Figure 6.30). PCR amplification of the F<sub>1</sub> plants (20 plants) with the *TaRAD51* gene primers (to amplify the ORF) as well as *AtRAD51* wild-type and T-DNA primers identified five different classes of F<sub>1</sub> progeny as against expected six different classes (Figure 6.29).

Three plants each containing the *CaMV35S::TaRAD51(D)* construct (expressing type) and without the *CaMV35S::TaRAD51(D)* construct (non-expressing type) were carried forward for an initial fluorescent analysis. The results of the analysis are presented as percentages of recombination. The preliminary results show that the percentages of recombination for the expressing type were two-fold higher than for the non-expressing type for all the plants analyzed. The average percentage of recombination of the expressing type was 38.7% compared to 17.8% of non-expressing type plants (Table 6.6). However this data is based on only few pollen grains (n=200) but needs to carry out at a large scale.

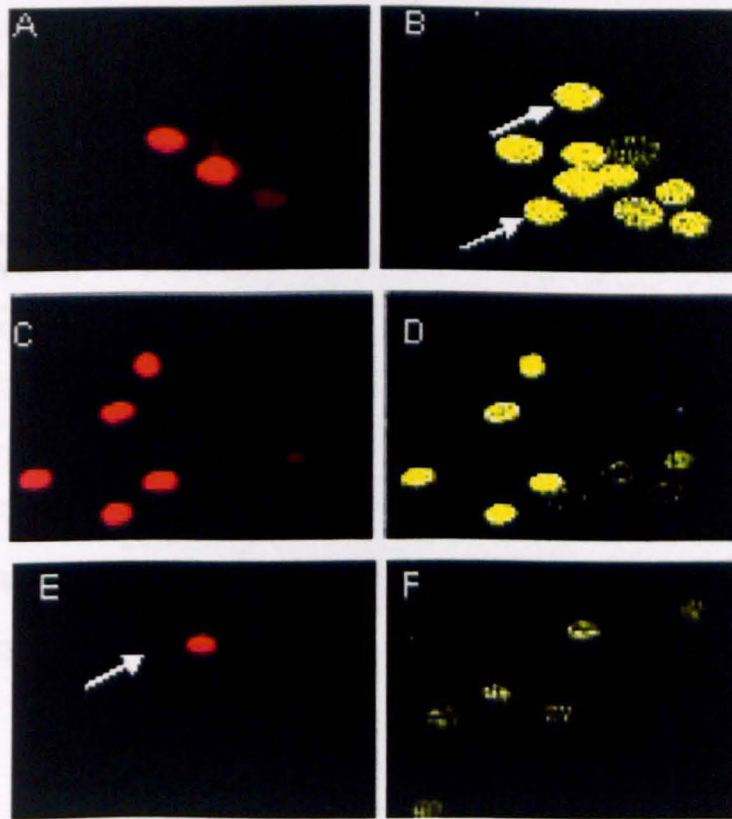




**Figure 6.28** Diagram depicting the crossing strategy employed to transfer the *TaRAD51* transgene from a heterozygous background into the fluorescent tetrad lines of Arabidopsis.



**Figure 6.29** PCR amplification of the F<sub>1</sub> progeny using the *TaRAD51* gene primers (to amplify the ORF) (a) and using *AtRAD51* WT and T-DNA primers (b); 1-20-F<sub>1</sub> progeny from the cross between *AtRAD51/atrad51+CaMV35S::TaRAD51(D)* and fluorescent tetrad line I3a, 21-water control and 22-positive control; M, 2-log ladder.



**Figure 6.30** Fluorescent analysis of *Arabidopsis* pollen. Only two markers were used to calculate crossover frequencies in both expressing and non-expressing plants. Three plants are shown here to indicate parental and recombinant pollen. AB, CD and EF are three plants visualized using two different filters (Red and Yellow). White arrows represent non-parental pollen and all the pollen without arrows are one of the parental type pollen. Recombinant pollen appears when a pollen grain expressed one FP, but not the other. No expression or expression of both FPs indicates the parental types.



**Table 6.6**    Scoring of pollen for the F<sub>1</sub> progeny of the cross between *AtRAD51/atrad51*+*TaRAD51(D)* and the fluorescent tetrad line I<sub>1</sub>a as recombination percentages.

Plant no.	Genotype of the plant for T-DNA insert	Genotype of the plant for <i>TaRAD51</i>	Pollen		Recombination (%)
			Total	Recombinants	
1	Heterozygous	Expressing	24	9	37.5
2	Heterozygous	Non-Expressing	31	9	21.0
4	Wild-type	Expressing	21	9	42.8
7	Heterozygous	Non-Expressing	14	3	21.4
8	Wild-type	Non-Expressing	18	2	11.1
10	Heterozygous	Expressing	45	16	35.5
		Expressing Average			38.7%
		Non-Expressing Average			17.8%

### 6.3 Discussion

#### 6.3.1 Wheat *RAD51* functional studies in Arabidopsis T-DNA lines

The three homoeologues of *TaRAD51* gene share 99% amino acid sequence similarity. Multiple alignments of the three *TaRAD51* homoeologues with the *AtRAD51* gene indicated a high percentage of similarity (81%) as well. Initially only one homoeologue (*D*) of *TaRAD51* was chosen for functional complementation studies in the present work. Comparative amino acid analysis and Predicted 3D overlays of *TaRAD51(D)* against *AtRAD51* revealed the highly conserved nature of these proteins.

The *AtRAD51* T-DNA insertional lines of Arabidopsis was obtained from Dr. Marie-Pascale Doutriaux (University Paris-Sud, France) and were characterized phenotypically and genotypically as a first step towards functional complementation assays with the *TaRAD51(D)* gene. No obvious differences were observed in all three (wild type, heterozygote and homozygous) self-fertilized segregants of *AtRAD51* heterozygote at any stage of vegetative development. At floral development, a reduction in filament length in homozygous mutants was observed. However, a reduction in pollen viability was observed between wild-type/heterozygote and homozygous mutants. Wild type/heterozygote plants of *AtRAD51* had completely viable pollen (98%) whereas the *atrads51* homozygous mutants had 0% pollen viability and was completely sterile. Expression evaluation of the *AtRAD51* gene confirms the results of Li *et al.*, (2004). Very high levels of expression of *AtRAD51* were observed in inflorescences of the wild-type plant indicating their active role during meiosis. The *AtRAD51* gene transcript was also found in leaf, root and silique of wild-type plants. No transcripts were found in any of the tissues analyzed in *Atrads51* homozygous T-DNA mutants indicated that the disruption of these genes by the T-DNA resulted in loss of function alleles and, presumably, no protein formation. It is possible that transcripts are produced, but are unable to accumulate, perhaps due to rapid degradation. This would also

be likely to produce a reduction (if not complete elimination) of the active protein for these genes.

Heterozygous mutants of the *AtRAD51* T-DNA lines were chosen for transformation because of complete sterility observed in *AtRAD51* homozygous mutant. It is still possible that *AtRAD51* mutants could have an effect on transformation efficiency, even in the heterozygous mutant state. An examination of pollen viability and seed set indicate that heterozygotes have wild-type levels of pollen viability and seed set which argues against a strong effect. Any increase in pollen viability or seed set would be an indication of a degree of functional complementation in homozygous mutant lines, when the wheat gene is over-expressed.

The transformant plants in the mutant background were first tested by RT-PCR to prove that only *TaRAD51* was transcribed in leaf, flower and silique, but that *AtRAD51* was not transcribed in the same tissues. Initial transformation results from the *CaMV35S::TaRAD51(D)* construct in a heterozygous T-DNA line of *AtRAD51* resulted in 12 T<sub>0</sub> transformants and genotypic analysis indicated that eight of them had only the wild-type allele while the remaining four had both a wild-type and a mutant alleles. No transformants with only mutant alleles were recovered. T<sub>1</sub> segregation analysis of one of the T<sub>0</sub> transformants indicated a segregation ratio of approximately 3:1 (31 transformants out of 43) ( $\chi^2 = 0.14$ ,  $P = 0.708$ ) carrying the transgene, indicating that the transgene is likely to have integrated as a single copy into the Arabidopsis genome. Complementation analysis of the *TaRAD51* transformants was calculated in terms of percentage of seed set compared to the non-transformed homozygous T-DNA mutant and wild-type. Preliminary results of the complementation analysis revealed that there is a significant increase in the pollen viability and a 18% and 0% seed set (compared to wild type) in the two different types of transformants (one completely sterile pollen and other partially viable pollen) over-expressing the *CaMV35S::TaRAD51(D)* construct. Across the two types of

*CaMV35S::TaRAD51(D)* construct plants there is 14% seed set with the above construct, compared with wild type plants. However, statistical analysis by Duncan revealed no significant difference in seed set between homozygous mutants and transformants with over-expressing *CaMV35S::TaRAD51(D)* construct.

Even though all the T<sub>1</sub> transformants were found to express transgene, some transformants were found to be completely sterile with regard to pollen viability and seed set and were phenotypically similar to the *Atrad51* homozygous mutant. The simplest explanation for potential differences between lines in terms of level of transgene expression would be that we expect the transgene to be segregating within this generation and individual plants may have one or two copies of the transgene. The low levels of seed set suggest that there is only borderline complementation of the mutation, so having two copies of the transgene rather than one could make a difference. This could be investigated in the next generation, where any heterozygotes would segregate for the Kanamycin resistance carried on the construct, while homozygotes for the construct would all be Kanamycin resistant. When analyzed by RT-PCR, the transgene seem to be expressed at similar levels in all the transformants. However, RT-PCR is not quantitative and Q-PCR could be used to further investigate this. Another possibility could be varying levels of translation of the transgene for which no experiments were done currently. As all plants are derived from a single transformation event, this also seems less likely.

To investigate the low levels of seed set in transformants and determine whether it might be due to lower pollen viability or due to the shorter anther filament structure in the transformants (also present in both T-DNA mutants), assisted self-pollination of the four siliques of the transformant with its own pollen and also crossing with wild-type pollen was performed. The results suggest that while assisted self-pollination does increase the size of the silique marginally compared to normal siliques on a wild-type plant, a more dramatic

effect is seen when out-crossing with wild-type pollen which leads to a clearly increased size of silique. However, the size of silique obtained with wild-type pollen is still not as great as a wild-type silique. This would imply that there is a still maternal defect because and the original T-DNA lines are reported as female sterile (Li *et al.*, 2004). Transformation with the over-expressing construct could not complement this defect fully. Based on these observations it may be concluded that the anther structure, the amount of fertile pollen (pollen viability) in the anthers as well as the fertility of stigma are important factors in the observed lack of substantial complementation. The fertility of the silique can be increased towards wild-type levels either by modifying the anther structure, effectively compensating for the short filament phenotype in the mutant or potentially by increasing the pollen available. The reason why some transformant plants display both sterile and fertile siliques on the same plant is intriguing and should be investigated further. The simplest explanation would be copy number of the transgene, but this is currently speculation.

The fact that *TaRAD51* could only complement the *Atrad51* mutation to a very limited extent may indicate that there might be a genuine functional difference between the *TaRAD51* and *AtRAD51* genes, perhaps due to the observed amino acid sequence differences at the N-terminal ends of these proteins which might affect how these proteins bind to DNA. However comparative amino acid alignments and predicted 3D modeling did not indicate functional disparity between the two proteins. The second possibility is that the *TaRAD51* gene may need a meiosis-specific or a strongly active promoter in meiosis (such as *AtRAD51* or *AtMRE11*) rather than the *CaMV35S* promoter. Li *et al.*, (2004) reported the successful complementation of the *Atrad51* mutant with a *CaMV35S* promoter construct where they used the *AtRAD51* gene. This would argue against this possibility, but it has not yet been formally tested within this work. If a larger sampling of To transformation events using a *CaMV35S*-driven *TaRAD51* failed to produce complementation, it would be argue that the *CaMV35S* promoter is not the intrinsic reason for lack of complete

complementation. Thirdly in the current work, the functional complementation was investigated using a cross species approach. Recent studies indicated only partial functional complementation between plant and yeast genes (Shaked *et al.* 2005; Klutstein *et al.*, 2008). Alternately, the wheat *RAD51* gene may not be a functional orthologue of the Arabidopsis *AtRAD51* gene and it could also be the case that the version chosen, while expressed in wheat, is not functionally active or as active as other homoeologues. The extremely limited seed set in the GFP fusion transformant (2%, compared with 14% for the non-fusion version) is a good argument that the fusion protein is less able to act as a substitute for the endogenous gene than the non-fusion wheat gene. This seems likely, as the protein is active as a nucleofilament and additional coding sequence of the GFP protein may well interfere with formation and biological activity of this nucleofilament.

The effects of over-expression of *CaMV35S::TaRAD51(D)* constructs on recombination frequencies and genetic distances were investigated using the Copenhagen fluorescent system in Arabidopsis. Preliminary results appear to show that the over-expression of *TaRAD51(D)* in Arabidopsis driven by a *CaMV35S* promoter produces an increase in genetic distances (2-fold; 38.7% recombination in wild-type segregants between YFP and RFP, compared with 17.8% in *TaRAD51* expressing plants). However, this needs to be further examined using larger numbers of pollen grains and also in a full *qrt* tetrad genetic background.

### 6.3.2 Wheat *DMC1* functional studies in Arabidopsis T-DNA lines

Multiple alignments of the amino acids of the three *TaDMC1* homoeologues against *AtDMC1* gene indicated a high percentage of similarity (87%). Because of high similarity of the three *TaDMC1* homoeologous gene sequences at the amino acid level (98%), initially only one homoeologue (*D*) of *TaDMC1* was chosen for functional complementation studies similar to the *TaRAD51(D)* gene

used in this study. Comparative amino acid analysis and Predicted 3D overlays of *TaDMC1(D)* against *AtDMC1* showed the highly conserved nature of these proteins at both amino acid level and structural level as well.

The *AtDMC1* T-DNA insertional lines of *Arabidopsis* obtained from Dr. Marie-Pascale Doutriaux (University Paris-Sud, France) were characterized phenotypically and genotypically. No phenotypic differences were observed in all three (wild type, heterozygote and homozygous) self-fertilized segregants of *AtDMC1* heterozygotes at any stage of vegetative plant development. However, at floral development a reduction in filament length in homozygous mutants was observed. During reproductive stages, a reduction in pollen viability was observed between wild-type/heterozygote and homozygous mutants. Wild type/heterozygote plants of *AtDMC1* were completely viable with 98% pollen viability whereas the *Atdmc1* homozygous mutant retained some pollen viability (17.5%). Because of this, *Atdmc1* homozygous mutants were able to produce limited numbers of seed (1.5%). Expression analysis by RT-PCR confirmed the observations reported by Couteau *et al.* (1999). Very high levels of expression of *AtDMC1* gene were observed in inflorescences of wild-type indicating their active role during meiosis. *AtDMC1* gene transcripts were also found in leaf, root and silique of wild-type plants. However no transcripts were found in any of the tissues analyzed in the *Atdmc1* homozygous T-DNA mutants indicating that the disruption of these genes by the T-DNA resulted in a loss of function allele and, presumably no protein formation. It is possible that transcripts are produced, but are unable to accumulate, due to rapid degradation. This would also be likely to produce a reduction (if not complete elimination) of active protein for these genes.

In this current work, the D homoeologue of the *TaDMC1* genes was cloned into over-expression constructs to test for potential complementation. Heterozygous mutants of the *AtDMC1* T-DNA lines were chosen for transformation because of very low fertility in *AtDMC1* mutants. The initial transformation results



using a *CaMV35S::TaDMC1(D)* over-expressing constructs indicated an initial recovery of three T<sub>0</sub> transformants but further T<sub>1</sub> segregation analysis could not recover any of these transformants. However, one transformant from *CaMV35S::TaDMC1(D)::GFP* was recovered in an Arabidopsis wild-type allele background. The *TaDMC1(D)::GFP* transformant in the wild-type background was crossed with the T-DNA *AtDMC1/Atdmc1* genotype and complementation will be studied in the F<sub>2</sub>. Due to insufficient time only the F<sub>1</sub> has so far been generated and the F<sub>2</sub> progeny are currently growing in the glasshouse.

In summary, in this current work multiple alignments of *TaRAD51* and *TaDMC1* homoeologous genes with *AtRAD51* and *AtDMC1* genes, respectively were done and revealed a high level of amino acid sequence similarity. An over-expression complementation construct involving the D homoeologue versions of *TaRAD51* and *TaDMC1* was prepared and used to transform homozygous *Atrad51* and *Atdmc1* mutants. Preliminary complementation results of the *TaRAD51* constructs without *GFP* fusion indicated a low level of complementation of the full T-DNA mutant for both pollen viability and seed set. No significant complementation was observed with *TaRAD51* constructs with *GFP* fusion. The complementation analysis of the *TaDMC1* has yet to be completed. Finally the over-expression of *TaRAD51(D)* in Arabidopsis appears to decrease crossover frequencies, based on an initial analysis using the Copenhagen fluorescent tetrad lines.

## CHAPTER 7: GENERAL DISCUSSION

There are three primary fates attained by the homoeologous genes during the course of evolution in allopolyploid species: functional diversification, gene silencing and retention of the original function (Lynch and Force, 2000). In polyploid species, knowledge of these processes will be essential for a better understanding of the role of a particular gene in question and the unraveling of associated crop traits. For example He *et al.*, (2003) estimated that 7–8% of genes are silenced in established wheat hexaploids and suggested that the genomic origin of silenced genes is non-random, and specifically that D homoeologue copies are silenced at a higher frequency than their A or B counterparts. In polyploid species, the function of a gene needs to be studied in terms of individual homoeologues rather than the gene orthologues as a whole, but very few studies to date have been aimed at studying the function of individual homoeologues for the genes involved in important processes in plant development (Ciaffi *et al.*, 2006).

Homologous Recombination is an essential process involved in maintaining genome integrity through accurate repair of DNA damage (DNA repair) and correct meiotic cellular division. HR has been extensively studied in many diverse eukaryotes, such as human, animal, fungi and plants owing to its importance. Homologous recombination involves tight regulation on many levels and many important and well-conserved components participate in these processes. This study has become more intense in recent years, since the realization that DNA repair genes (Rad51 and Dmc1) are found to interact with Brca2, the mutation of which causes breast cancer in humans (Dray *et al.*, 2006). During meiosis, HR is the process underlying strand exchange between paternal and maternal chromosomes and thereby generates diversity among the progeny derived from them (reviewed by Filippo *et al.*, 2008). The two major strand-exchange genes involved in HR in eukaryotes are *RAD51* and *DMC1*. The Rad51 recombinase is mainly involved in strand exchange of DNA

between homologous chromosomes and is active in mitosis and meiosis. The Dmc1 recombinase promotes similar associations between homologous DNA but is active in meiosis only (Masson and West, 2001). Because it has been demonstrated that *RAD51* and *DMC1* gene orthologues are involved in HR in other cereal species and Arabidopsis (model species) (Doutriaux *et al.*, 1998; Franklin *et al.*, 1999; Kathiseran *et al.*, 2002; Li *et al.*, 2007) these genes have been investigated for their role in HR in hexaploid wheat. The study of meiotic cellular processes and HR genes involved in meiosis are important as current breeding strategies for hexaploid wheat have been hampered by the relatively low frequencies and strongly localized distribution of recombination events during meiosis. Broadening our current knowledge of the molecular basis of meiosis with respect to the functioning of HR genes would not only enable further understanding of meiosis in hexaploid wheat but also present opportunities for wheat breeding programs to exploit increases in recombination frequencies which would help to reduce linkage drag during gene introgression and also increase the number of breakpoints in meiosis. The roles of HR genes, *RAD51* and *DMC1* in wheat have been investigated in this current thesis. Very few meiotic genes have been characterized to date in hexaploid wheat, partly due to the large genome size and complexity of this polyploid.

Before cloning the HR cDNA homoeologues from hexaploid wheat, our initial focus was to analyse microarray transcript profiling of both *RAD51* and *DMC1* genes in Arabidopsis, rice and wheat to confirm the results of previous reports and as a starting point for planning experiments (Chapter 3). For all the three cDNA homoeologues of *TaRAD51* and *TaDMC1* only one probe set could be retrieved even though the E-values of the probe set differed with the homoeologues. This is probably because the Affymetrix GeneChip does not allow for the differentiation of individual homoeologues. The results of the microarray study using available microarray databases indicated strong similarity between the expression patterns of the *TaRAD51* and corresponding

orthologues in rice (*OsRAD51A1* and *OsRAD51A2*) and Arabidopsis (*AtRAD51*). The *RAD51* gene was found to expressed in both mitotically and meiotically active tissues in all three plants. This is similar to RT and Q-PCR results reported by Ding *et al.*, (2001), Li *et al.*, (2004) and Khoo *et al.*, (2008). But the expression of *TaRAD51* in mitotically inactive tissue such as mature leaves is intriguing and a likely explanation would be that the low level expression observed us due to its role in DNA repair.

The microarray analysis for *TaDMC1* indicated that it is expressed in both mitotically and meiotically active tissues and follows a similar expression profile pattern as the Arabidopsis and rice *DMC1* genes. This is consistent with Doutriaux *et al.*, (1998); Kathiseran *et al.*, (2002) and Khoo *et al.*, (2008) using RT and Q-PCR techniques. Unlike in yeast, *DMC1* in Arabidopsis, rice and wheat is expressed in both vegetative and reproductive tissues. In fact the highest expression was found to be in cell suspension cultures. URS1 was found to be the main regulator for this kind of expression and was first observed in yeast (Vershon *et al.*, 1992; Prinz *et al.*, 1995). URS1 is a transcriptional repressor site found in the promoters of a wide variety of yeast genes. During vegetative growth in yeast, transcriptional activation of the meiotic promoter is repressed by the factors that act at the URS1 site. However, after cells enter meiosis, URS1 switches its function and becomes an activator site and results in high level of expression of meiotic gene. In the context of meiotic promoters, URS1 sites act as repressor sequences during mitosis and function as activator sites during meiosis. Doutriaux *et al.* (1998) was of the opinion that loss of this negative promoter element was responsible for non-meiotic repression and could explain the observed regulated mitotic expression of *AtDMC1* in cultured cells. The detection of *AtDMC1* transcripts in root and leaf as well suggests that this regulation may not simply be restricted to the cell suspension cultures.

At the beginning of the studies in this thesis, genetic analysis of the *RAD51* and *DMC1* genes and their homoeologues in hexaploid wheat had not been reported. To gain more insight into the HR pathway in hexaploid wheat and compare it with rice and Arabidopsis, the *RAD51* and *DMC1* cDNA homoeologues were cloned from hexaploid wheat (Chapter 4). A remarkable level of synteny has been observed between the genomes of cereals especially wheat and rice (Kurata *et al.*, 1994) and this has allowed the cloning of the wheat *RAD51* cDNA sequence using a rice *RAD51A1* cDNA sequence as a starting point. A partial cDNA sequence of wheat *DMC1* was reported in the database and was used as the basis to clone the full-length wheat *DMC1* cDNA sequence. The individual cDNA homoeologues of *TaRAD51* and *TaDMC1* were cloned based on the sequence differences identified among the coding and non-coding regions of these cDNA sequences. Genome-specific primers were designed based on sequence differences identified and a PCR-based genome-specific test devised. A similar strategy was used by Blake *et al* (2004) to develop genome-specific primer sets and isolated the starch biosynthesis-related genes from the three genomes of hexaploid wheat. Recently Khoo *et al.*, (2008) reported coding sequences of *TaRAD51* and *TaDMC1* and preliminary molecular characterization of the *TaRAD51* and *TaDMC1* cDNA sequences, but they do not report the individual homoeologues.

This study has found that there is only one *TaRAD51* gene per haploid genome. However, recently Khoo *et al.*, (2008) reported that there are two paralogues, *TaRAD51A1* and *TaRAD51A2*, both located on the same chromosome Group 7. This result is in contrast to other reports of *RAD51* gene in rice and maize where two paralogues of *RAD51* were found to be located on different chromosomes (Franklin *et al.*, 1999 and Li *et al.*, 2007) and are almost certainly the result of duplication between Chr.11 and Chr.12, in the case of rice (Kathiseran *et al.*, 2002). Multiple alignments of the *TaRAD51* homoeologues cloned in this work with the published *TaRAD51A1* and *TaRAD51A2* revealed that *TaRAD51A1* was found to be similar to the *TaRAD51(D)* homoeologue and

*TaRAD51A2* was found to be a truncated homoeologue of *TaRAD51(A)*. It may be that in the cultivar investigated by Khoo *et al.*, (2008) this is due to a gene translocation/rearrangement event. We found one *TaDMC1* per genome and were able to identify the reported *TaDMC1* sequence by Khoo *et al.*, (2008) to be the (D) homoeologue of *TaDMC1*.

In this study, the three cDNA homoeologues of *TaRAD51* and three cDNA homoeologues of *TaDMC1* were found to share high amino acid identity of 99% and 98%, respectively, albeit with few amino acid substitutions between the three cDNA homoeologues of respective genes. SIFT analysis revealed that the amino acid substitutions between the three cDNA homoeologues of *TaRAD51* and *TaDMC1* genes are all tolerable and were likely to have no deleterious effects on the structure of final protein. The amino acid analysis of the three cDNA homoeologues of *TaRAD51* and the three cDNA homoeologues of *TaDMC1* homoeologues revealed high levels of identity to the rice and Arabidopsis orthologues. Comparative amino acid sequence analysis and 3D modeling revealed high levels of structural conservation among these proteins in Arabidopsis, rice and wheat. To further elucidate the evolutionary relationships within the family of *RAD51* and *DMC1* paralogues, phylogenetic analysis was performed. The analysis revealed that all three cDNA homoeologues of *TaRAD51* and all three cDNA homoeologues of *TaDMC1* were closely related to each other, as would be expected based on their amino acid sequences. The analysis also revealed that there is a clear separation between Rad51, Dmc1 and the other Rad51 protein paralogues. The phylogenetic tree created in this research confirmed the reports of previous phylogenetic trees from the literature and reviews (Bleuyard *et al.*, 2005; Lin *et al.*, 2006; Li *et al.*, 2007 and Khoo *et al.*, 2008).

The analysis of expression of individual cDNA homoeologues of *TaRAD51* and *TaDMC1* was performed using Q-PCR in this study. Previously Northern blotting and cDNA-SSCP (Adams and Cronn, 2003) were used for expression

analysis of cDNA homoeologues but they have certain limitations, such as use of radioactive materials or the use of fluorescence to measure gene expression. Q-PCR has been used in various applications for expression analysis and in the current research, the same technique was used successfully. The Q-PCR results in this research revealed that all the three cDNA homoeologues of *TaRAD51* and *TaDMC1* were expressed in all stages of meiosis and the expression patterns of the three cDNA homoeologues of *TaRAD51* were similar, as were three cDNA homoeologues of *TaDMC1*. Considering that the three homoeologues of many wheat genes are frequently all expressed (70%, as estimated by Mochida *et al.*, 2004) it is not surprising to find the expression of all three homoeologues of *TaRAD51* and *TaDMC1*. This result also agrees with the Q-PCR result reported by Khoo *et al.*, (2008) for the expression of the (D) homoeologue of *TaRAD51* and *TaDMC1*. The expression level of co-expressed homoeologues can sometimes be equal (Wknox-1, Morimoto *et al.*, 2005; Wsep, Shitsukawa *et al.*, 2007) or vary (TaBx, Nomura *et al.*, 2005; TaGA20ox1, Appleford *et al.*, 2006). In the current study, the expression of one of the cDNA homoeologues (*TaRAD51(B)* and *TaDMC1(B)*) was higher when compared to the other two homoeologues, which may indicate an uneven contribution by the different homoeologues to meiotic recombination.. The Q-PCR results obtained here were shown to be consistent with the microarray expression analysis undertaken earlier.

In this research, two broad experimental approaches were employed for genetic analysis of *TaRAD51* and *TaDMC1* homoeologues. One is '*in planta*' analysis using Forward and Reverse Genetics in hexaploid wheat (Chapter 5). Forward and reverse genetics strategies have been successfully employed to study many genes and their homoeologues in hexaploid wheat (Slade *et al.*, 2005). Therefore, these strategies were used to try to study the function of *RAD51* and *DMC1* gene homoeologues in hexaploid wheat. The main aim of Forward genetics is to perform phenotypic screening of the mutagenized population for any observable characters leading to ear sterility. As both *RAD51* and *DMC1*



genes are involved in meiosis, any disruption of the expression of these genes in hexaploid wheat could result in partial or complete fertility depending on the level of mutation and number of homoeologues deleted in the population. Reverse genetic analysis was used to perform genotypic screening of the mutagenized population for the identification of lines with deletions in target genes and their homoeologues (*RAD51* and *DMC1*) and to try to assign their function based on phenotype. The second method is a cross species approach which utilizes *Arabidopsis* T-DNA lines to study the function of the target genes (*TaRAD51* and *TaDMC1* gene homoeologues) by trying to complement the mutant phenotype (Chapter 6). This is an indirect method of analyzing gene function because it is performed in a heterologous species.

In general it is more difficult to use Forward Genetic strategies in bread wheat because the phenotype of the mutant plant with deletions in the target gene might be masked as bread wheat genome could be highly buffered because of polyploidy. Reverse genetics screening has been successfully applied to identify deletions in the targeted genes of many plant species (McKinney *et al.*, 1995; Krysan *et al.*, 1996). For both Forward and Reverse genetics, the use of an appropriate method to induce mutations in the population is important for polyploid species. Because methods such as chemical mutagenesis, gene silencing by antisense/sense suppression and intron-spliced hairpin often produce partial inactivation of the endogenous gene it is not possible to predict the extent of targeted gene disruption, data interpretation is difficult (Höfgen *et al.*, 1994; van der Krol *et al.*, 1990). Therefore the use of wheat populations irradiated with Fast neutron and Gamma rays which produce bigger deletions compared to the above methods were utilized to study the role of target genes and their homoeologues (*TaRAD51* and *TaDMC1*).

Forward genetic screening of a Highbury Fast neutron population indicated the absence of any detectable deletion with phenotyping of the lines. One possible reason could be that the detection method used for screening the frequency of

deletions left in the population. This is because only ten microsatellite markers were used for initial screening. Secondly the characters that we are looking at in this generation (partial sterility/partial fertility) could have been selected against in previous generations. Although morphological screening yielded four potential lines with reduced fertility among 1060M<sub>8</sub> lines, further analysis indicated lack of any deletions in the target genes. The observed reduced fertility in those four lines may be because of deletions in some other genes mostly related to meiosis other than target genes or even due to non-meiotic, energy production (as pollen production is the most energy intensive process a plant undergoes or environmental factors. Reverse Genetic screening of the 200 lines of the Fast neutron population with ten microsatellite markers spread across the whole genome is consistent with Forward genetic screening result and no deletions were observed with any of the markers indicative of the fact that the population has little or no deletions left in them.

Unlike for the Highbury fast neutron mutant population, we were able to demonstrate that deletion mutants can be identified for targeted plant genes and homoeologues (*TaRAD51* and *TaDMC1*) by screening of the Paragon Gamma mutant population using a simple PCR screening procedure. Li *et al.* (2001) showed the effectiveness of similar PCR-based reverse genetic approach for detecting mid-sized deletions (1 kbp) caused by fast-neutron irradiation and was able to find deletion mutants for more than 80% of the 25 loci tested in a population of 51, 840 Arabidopsis fast neutron lines. The isolation of deletion mutants in target genes in the Paragon population could support our earlier conclusions that the Highbury Fast neutron population is at very advance stage (M<sub>8</sub>) compared to the Paragon gamma mutant population (M<sub>3</sub>) by which time most of the mutations either become homozygous or have been selected against in the previous generations. Although wheat appears to be particularly suitable for mutant screening experiments since it is able to carry a very high mutation load, only one deletion line of *TaRAD51(B)*, one deletion line of *TaDMC1(A)* and one deletion line of *TaDMC1(D)* survived, despite the initial detection of

at least one line for each homoeologue at the  $M_2$  stage. Most of the gamma ray-induced mutations create large-scale deletions of up to 6 Mb (Naito *et al.*, 2005) and in this context lower rate of survival could be because of such large deletions. The  $M_3$  progeny from the deletion lines grown in Nottingham revealed that the deletion for the gene has been inherited to the  $M_3$  generation and RT-PCR expression analysis confirmed that the deletion lines did not accumulate transcripts of the relevant genes.

Preliminary phenotypic analysis of one deletion line (*TaRAD51(B)*) and two deletions lines (*TaDMC1(A)* and *TaDMC1(D)*) indicated that both *TaRAD51* and *TaDMC1* genes are dispensable for vegetative growth of the plants. This observation is consistent with the earlier reports of these genes in *Arabidopsis* where a single copy of *AtRAD51* and *AtDMC1* exists (Doutriaux *et al.*, 1998; Li *et al.*, 2004). In plants where there are two copies of the *RAD51* gene exists such as *Zmrad51a* and *Zmrad51b* in maize, the plants homozygous for single mutants of either *Zmrad51a* or *Zmrad51b* develop normally (Franklin *et al.*, 2003). Based on the above results it is not surprising that single homoeologous gene mutants are viable vegetatively in hexaploid wheat. However, to be able to investigate any effect of the deletion of *TaRAD51* and *TaDMC1* homoeologues on vegetative growth of the wheat plant, the development of double or triple deletions would be necessary, due to potential functional redundancy in hexaploid wheat. As these genes are involved in meiotic recombination, the possible effect of deletions on processes such as meiosis and the consequent level of pollen viability were examined. The results showed that there are significant differences between the pollen viability of unmutated Paragon and that of the deletion lines. This implies that there might be abnormalities in meiosis occurring in these deletion lines. The cytological examination of male meiosis performed in these deletion lines indicated contrasting results for the *TaRAD51(B)* and for *TaDMC1(A)* and *TaDMC1(D)* deletions lines when comparing pollen viability results. Even though a pollen viability of only 70% was obtained for the *TaRAD51(B)* deletion line when

compared to unmutated paragon (99%), no meiotic abnormalities were found in any of these lines. This may suggest that the observed low pollen viability has been caused by a factor other than the effect of the deletion of the *TaRAD51(B)* homoeologue on meiotic progression. For the *TaDMC1(A)* and *TaDMC1(D)* deletion lines, meiotic abnormalities were observed during diakinesis and tetrad formation (frequent dyads are observed among tetrads). So the observed pollen viability of 67% and 54%, respectively, for *TaDMC1(A)* and *TaDMC1(D)* deletion lines when compared to unmutated paragon (99%) may be at least partly attributed to defects in meiosis. Literature searches on the deletion of *RAD51* and *DMC1* in other species such as rice, maize and *Arabidopsis* indicated very low pollen viability (0%-10%) (Doutriaux *et al.*, 1998; Franklin *et al.*, 1999; Kathiseran *et al.*, 2002; Li *et al.*, 2007) accompanied with defects in meiosis (SC assembly and chromosomal instability and fragmentation in *Atrad51* mutants). In the nullisomics of group 5, frequent meiotic aberrations have been observed during diakinesis and resulting univalents are constantly seen during Prophase I (Sears, 1954).

Ear fertility (grain/spikelet) of the deletion lines compared with unmutated Paragon indicated that the combined effects of paternal and (presumably) maternal defects led to the ear fertility dropping to 93%, 54% and 18% of wild-type levels in the *TaRAD51(B)*, *TaDMC1(A)* and *TaDMC1(D)* deletion lines, respectively. The ear fertility results of the *TaRAD51(B)* deletion line indicated only a 7% reduction in ear fertility when compared to wild-type, suggesting that this highly expressed homoeologue can be missing from meiosis with minimal effects on the ear fertility. This is in contrast to *Atrad51* but similar to the *Zmrad51a1* and *Zmrad51a2* mutants. The *Atrad51* mutant was completely sterile but *Zmrad51a1* and *Zmrad51a2* mutants were completely fertile but the double mutants are completely sterile. This implies that the deletion of *TaRAD51(B)* homoeologue has minimal effects similar to the situation in maize (however with reduced yield compared to wild-type, so not identical to maize). If the other two homoeologous mutants have similar effects on ear fertility and

the triple deletions have complete sterility then it is appropriate to say that polyploid wheat behaves similar to diploid maize, which is an ancient tetraploid having two copies of the genes. The other possibility is even though the *TaRAD51(B)* homoeologue is highly expressed in meiosis it might be a non-functional homoeologue. However comparative amino acid analysis and SIFT analysis does not give any evidence to indicate a non-functional homoeologue. Assuming that *TaRAD51(B)* is functional it suggests a lot of buffering from the other two homoeologues of *TaRAD51*.

The aneuploids of bread wheat, such as the Nullisomic, lines (Sears 1954, 1966; Sears and Sears 1978) have been extensively used to assign genes and molecular markers to individual chromosomes and chromosome arms. Sears (1954) observed that the three Nullisomics belonging to the group 7 chromosomes have the least effect upon the phenotype of the euploid when compared with Nullisomics for chromosomes of other homoeologous groups. Based on the Nullisomic observation it is not surprising the deletion of *TaRAD51* has least effect on the ear fertility of the lines investigated. Until deletion lines for the other two *TaRAD51* homoeologues (*A* and *D*) can be recovered, characterized and combined, the current work remains inconclusive assumption. The drastic reduction in ear fertility of *TaDMC1(D)* and *TaDMC1(A)* homoeologue mutants (18% and 54% of wild-type levels) was observed suggesting that they play a significant role in meiosis. Previously Q-PCR results showed relatively higher levels of expression for *B* and *D* homoeologues compared to *A* homoeologue and by comparing the ear fertility scores, the order of the relative importance of these homoeologues would be expected to be  $D > B > A$ , assuming similar levels of protein activity for each homoeologue Unless the deletion lines for the *B* homoeologue can be recovered this remains speculation. Sear (1954) reported that Nullisomics of group 5 have narrow leaves and slender culms, are late in maturity, spikes are reduced in size and the have small glumes and seeds. Sears and Sears (1978) also mentioned that important genes related to fertility are located on these chromosomes.

Literature searches on the genes present on the group 5 chromosomes revealed a number of genes related to abiotic stress and the current *TaDMC1* gene. As complete sterility is not observed in the *TaDMC1* gene homoeologous mutants, it is likely that there is buffering effect of the other homoeologous genes for this gene. Unless the deletion lines for the *TaDMC1* homoeologue (*B*) can be recovered and characterized this remains unproven. Future research will also be required to further investigate these effects in wheat by the development of double and triple deletion lines involving *TaDMC1* gene homoeologues.

The Forward and Reverse genetic approaches illustrated that these techniques can be used to characterize the effects of deletion of individual homoeologue of wheat genes, but where there is not complete functional redundancy/compensation between homeologues, some deletion lines may be difficult to obtain. When a phenotype is observed in a deletion line using Forward and Reverse Genetic approaches, complementation tests (by transforming the mutant with the wild type gene) need to be conducted to link gene to phenotype. Because complementation tests are difficult to conduct in hexaploid wheat, in the present research Arabidopsis T-DNA mutant lines with insertions in target genes (*RAD51* and *DMC1*) were used to test for complementation by the wheat target genes. Arabidopsis T-DNA lines were used to examine the effects of over-expression of *TaRAD51* and *TaDMC1* genes on the functional complementation of the appropriate mutant line and on recombination frequencies in Arabidopsis (Chapter 6). Apart from being a relatively quick and *in planta* analysis, this system also adds additional data to the findings of genetic analysis obtained through Forward and Reverse Genetics screening. More importantly this experiment illustrates the potential of Arabidopsis T-DNA lines for functional complementation of heterologous genes in Arabidopsis. Only one homoeologue of the *TaRAD51* and *TaDMC1* genes was tested for functional complementation studies because of highly similar sequence of the homoeologous sequences. There could still be subtle differences between the homoeologues and this needs to be considered, even if

the sequence and SIFT analysis suggested that no major functional differences were expected between the homoeologues.

Initial results of the functional complementation tests indicated that overexpressing *CaMV35S::TaRAD51(D)* constructs without the GFP reporter gene gave a low level of functional complementation in the homozygous mutant plant background. On the other hand over-expressing *CaMV35S::TaRAD51(D)* constructs with the GFP reporter gene gave no significant functional complementation in the homozygous mutant plant background. In both cases, the morphology of the siliques were quite similar to the homozygous mutants, although the pollen viability was slightly increased (non-significant). However the expressing lines with the *CaMV35S::TaRAD51(D)* constructs without GFP reporter constructs produced 14% seed compared to 0% in the homozygous mutant (compared to the wild type controls). On the hand, the expressing lines with the *CaMV35S::TaRAD51(D)* construct with a GFP reporter fusion only produced 2% seed compared to 0% in the homozygous mutant. This implies that although there is an increase in pollen viability and the gene was expressed, the *TaRAD51*-overexpressing plants did not show strong complementation of mutant phenotype. This observation is different from results obtained from *AtRAD51*-overexpression Arabidopsis. Li *et al.*, (2004) reported that *AtRAD51* overexpression under a *CaMV35S* promoter in Arabidopsis was able to complement the mutant phenotype fully. Potentially, this might argue that the wheat *TaRAD51* gene may not be a complete functional orthologue of the Arabidopsis *AtRAD51* gene since their proteins shared only about 81% identity and that functional dissimilarities between the wheat and Arabidopsis *RAD51* genes did not allow complete functional complementation, even if the level of expression in meiosis of the transgene driven by *CaMV35S* was sufficient. The very low functional complementation of the *CaMV35S::TaRAD51(D)* construct with the GFP reporter gene compared to the construct without GFP could be explained on the basis that fusing the GFP fragment (246 amino acids) onto the



gene may interfere with protein-protein interaction and the establishment of the nucleofilament.

The lack of elongation of siliques in the homozygous mutants was thought to be due to the short anther filament phenotype and the low yield of pollen (Li *et al.*, 2004). The anthers of *TaRAD51*-overexpressing plants have shortened filaments similar to *atrad51* homozygous mutant plants. However an increase in pollen yield and viability was observed. Elongated siliques compared to normal silique were obtained by assisted self-pollination and even more elongated siliques were obtained when assisted crossing used wild-type pollen. Unlike the limited elongation of siliques with assisted self-pollination, the later ones produced near normal siliques. However this data needs to be considered carefully because this is based on a limited number of crossings. Over-expressing constructs of *TaDMC1* have been produced in a heterozygous *AtDMC1/Atdmc1* background, but lack of time prevented the further functional complementation analysis in an *Atdmc1* homozygous mutant background.

*TaRAD51*-overexpressing plants in an *Atrad51* homozygous background were crossed to one of the Copenhagen fluorescent tetrad lines to investigate if there are changes in recombination frequencies and genetic distances associated with expression of the construct, compared to non-expressing plants. Looking at a genetic interval on Chromosomes 3, we observed an average 46% increase in genetic distances in meiotic crossing over in the *TaRAD51*-overexpressing plants, compared to non-expressing segregants. This is similar to the observation by Betzner and Stefan *et al.*, (WO/2002/008432) where the over-expressing constructs using the *AtRAD51* gene in Arabidopsis was found to increase the genetic distance by 2-fold. Berchowitz *et al.*, (2007) reported that the *Atmus81* mutants showed a 9% reduction in the interval on Chromosome 3 (6,710 tetrads) using this technique. The reduction in genetic distances was observed in *Atmus81* mutants related to meiosis where a lack of meiotic crossovers leads to decreased genetic distance. This provisional result was

based on using a non-quartet background and was also based on very limited numbers of pollen grains ( $n=200$ ). So, the results of this assay needs to be considered with caution, until a more extensive analysis can be completed.

The prospects for further research to characterize the HR genes and their homoeologues in hexaploid wheat depends on alternate strategies from using mutagenized populations by forward and reverse genetic approaches. Firstly, it is both time consuming and laborious to screen large populations for mutants. Secondly even though deletion mutants screened from mutagenized population are an excellent source of material for analyzing the function of the target genes there are a few potential limitations. The observed phenotype could be caused by deletion in some other gene apart from the target gene. For example *Arabidopsis* meiotic genes such as *RAD51*, *MND1*, *MEI1*, and *SPO11* exhibit decreased pollen viability when disrupted, either as a result of chromosome fragmentation or segregation defects and ultimately this leads to sterility (Li *et al.*, 2004; Kerzendorfer *et al.*, 2006; Grelon *et al.*, 2003). In such cases it is important to backcross the deletion mutant with original parent line from which it is derived to remove any additional deletions. The deletion lines for the target genes (*TaRAD51* and *TaDMC1*) in this study have only been studied for one generation and so to relate the observed phenotype to the deletions in target genes only and to remove any additional deletions, future research should aim to backcross the mutants with the parent.

The low level of polymorphism in wheat due to its polyploidy and the recent origin of the hexaploid wheat genome have hampered both forward genetics and the genetic mapping of wheat. To overcome this, Endo and Gill (1996) have proposed the use of deletion lines. They have reported the isolation of 436 deletions which have been extensively used in molecular mapping of the wheat genome. Apart from genetic mapping they can be used for forward genetic approaches also. They are easy to maintain unlike nullisomics and the gene can be mapped on the deletion lines easily.

Alternately in the absence of full set of deletion lines for the target genes and their homoeologues, the availability of characterized Nulli-tetrasomic lines gives us an opportunity to substitute for the deletion lines. In addition, they can also be used to estimate the relative contribution each genome makes to the full transcription profile and ask whether the gene product of one genome is more frequently represented than expected. CS Nulli-tetrasomic cytogenetic stocks were used in this research to locate the homoeologues and confirm their genome specificity. Nulli-tetrasomic analysis revealed that the *TaRAD51* and *TaDMC1* genes were located on Group 7 and 5 chromosomes, respectively, with a single copy on each of the genome (A, B and D).

Finally future studies should focus on the mechanism by which these HR genes are regulated in hexaploid wheat, and of particular interest is how these individual homoeologues are involved in different stages of meiotic recombination, particularly homology search, as well as crossing-over. Homologous recombination is a fundamental process in maintaining cell integrity and also in evolution, through meiotic recombination and the reassortment of genes on individual chromosomes. To be able to manipulate this process could have major implications for conventional breeding and also for gene targeting for transgenic crop improvement approaches. This study is the first to investigate the function of the two major strand-exchange genes in hexaploid wheat, one of the two major cereal crops for world food production.

## **7.1 Future perspectives**

### **7.1.1 Isolation of full-length genomic clones of *TaRAD51* and *TaDMC1* homoeologous genes**

The full-length genomic clones of *TaRAD51-7A*, *TaDMC1-5B* and *TaDMC1-5D* have already been isolated from hexaploid wheat in the lab and the isolation of the rest of the genomic clones: *TaRAD51-7B*, *TaRAD51-7D* and *TaDMC1-5A* is in progress. The aim of this research is to elucidate the structure of

genomic clones for intron/exon organization, differences of intronic regions among the three homoeologous genes of *TaRAD51* and also among the three homoeologous genes of *TaDMC1*. This background work helps us to make a comparison between the genomic organization of wheat *RAD51* and *DMC1* genes with that of other plants.

### **7.1.2 Screening for the isolation of *TaRAD51* (A), (D) and *TaDMC1* (B) homoeologous gene mutants by Reverse Genetic analysis**

As only a few deletions of the homoeologous gene mutants have been obtained for the current work, future work aims to screen an enlarged mutagenized population of Paragon wheat for the isolation of all the homoeologue deletions of *TaRAD51* and *TaDMC1*. This would serve better for comparisons of expression analysis and phenotypic analysis. Also they would be useful for the production of double or triple deletion lines involving the individual deletion lines.

### **7.1.3 Over-expression studies of the *TaDMC1* gene in *Atdmc1* homozygous mutants for possible complementation**

The current work aimed to carry out over-expression studies of *TaDMC1* constructs in an *Atdmc1* homozygous mutant background for possible complementation. One transformant with the CaMV35S::*TaDMC1*(D):GFP construct was obtained in a wild-type background. In order to transfer the construct into *Atdmc1* homozygous background, the transformant was crossed with a heterozygous *AtDMC1/Atdmc1* background and a F<sub>1</sub> heterozygote *AtDMC1/Atdmc1* with the over-expressing *TaDMC1* constructs was obtained. Future work aims to segregate this heterozygous *AtDMC1/Atdmc1* plants to analyze homozygous *Atdmc1/Atdmc1* plants with over-expressing *TaDMC1* constructs for possible complementation of mutant phenotype. Given the reduction in complementation observed in the seed set for *TaRAD51*(D)::GFP compared with the GFP- version, a similar approach with *TaDMC1*(D) is also needed.

#### **7.1.4 Use of over-expressing constructs with gene specific (*RAD51/DMC1*) promoters**

The over-expressing *CaMV35S::TaRAD51(D)* construct could not complement the mutant phenotype of the *Atrad51* homozygous plants. In order to exclude a role of the promoter in this limited complementation, gene-specific promoters (*AtRAD51/AtDMC1*) will be evaluated for their ability to drive the *TaRAD51* and *TaDMC1* genes and produce higher levels of complementation in the appropriate mutant background. Failure to achieve full complementation of the appropriate mutant phenotype would argue for functional differences between the wheat and Arabidopsis versions of these genes.

#### **7.1.5 Studying the recombination frequency in over-expressing lines of *TaDMC1***

Only over-expressing lines of *TaRAD51* have been used to study recombination frequencies and genetic distances by crossing with the fluorescent tetrad lines so far. Future research aims to also evaluate over-expressing lines of *TaDMC1* for possible effects on recombination frequencies and genetic distances by crossing with appropriate fluorescent tetrad lines of Arabidopsis.

## References

- Aboussekhra, A., Chanet, R., Adjiri, A. and Fabre, F (1992) Semi-dominant suppressors of Srs2 helicase mutations of *Saccharomyces cerevisiae* map in the *RAD51* gene, whose sequence predicts a protein with similarities to prokaryotic RecA proteins. *Mol. Cell. Biol.* 12, 3224-3234
- Adams, K.L., R.C. Cronn, R. Percifield, and J.F. Wendel (2003) Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Sciences USA* 100, 4649-4654
- Adams, K. L., R. Cronn, R. Percifield and J. F. Wendel (2004) Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. *Genetics* 168, 2217-2226
- Akhunov ED, Akhunova AR, Linkiewicz AM, Dubcovsky J, Hummel D, Lazo G, Chao S, Anderson OD, David J, Qi L, Echalié B, Gill BS, Miftahudin, Gustafson JP, La Rota M, Sorrells ME, Zhang D, Nguyen HT, Kalavacharla V, Hossain K, Kianian SF, Peng J, Lapitan NL, Wennerlind EJ, Nduati V, Anderson JA, Sidhu D, Gill KS, McGuire PE, Qualset CO, Dvorak J (2003) Syteny perturbations between wheat homoeologous chromosomes by locus duplications and deletions correlate with recombination rates along chromosomes arms. *Proc Natl Acad Sci USA* 100, 10836-10841
- Albala, J.S., Thelen, M.P, M Prange, C, M Fan, W., Christensen, M, M Thompson, L.H, and Lennon, G, G (1997) Identification of a novel human RAD51 homolog, RAD51B. *Genomics* 46, 476-479
- Alexander, M.P (1969) Differential staining of aborted and non aborted pollen. *Stain Technology* 44, 117-122

- Al-Kaff Nadia, Emilie Knight , Isabelle Bertin , Tracie Foote , Nicola Hart , Simon Griffiths , and Graham Moore (2008) Detailed Dissection of the Chromosomal Region Containing the *Ph1* Locus in Wheat *Triticum aestivum*: With Deletion Mutants and Expression Profiling, *Ann Bot* 101, 863-872
- Alkharouf Nadim W., and Benjamin F. Matthews (2004) SGMD: the Soybean Genomics and Microarray Database *Nucl. Acids Res.* 32, 398-400
- Alonso, J.M. and J.R. Ecker (2006) Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in Arabidopsis. *Nat. Rev. Genet* 7, 524
- Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. and Ecker, J.R (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301, 653-657
- Amundsen SK, Smith GR (2003) Interchangeable parts of the *Escherichia coli* recombination machinery. *Cell* 112, 741-744
- Anderson, L.K., Offenberg, H.H., Verkuijlen, W.M.H.C., and Heyting, C. (1997) RecA-like proteins are components of early meiotic nodules in lily. *Proc. Natl. Acad. Sci. USA* 94, 6868-6873
- Andreas W Schreiber, Tim Sutton, Rico A Caldo, Elena Kalashyan, Ben Lovell, Gwenda Mayo, Gary J Muehlbauer, Arnis Druka, Robbie Waugh, Roger P Wise, Peter Langridge and Ute Baumann (2009) Comparative transcriptomics in the Triticeae. *BMC Genomics* 10:285
- Appleford NE, Evans DJ, Lenton JR, Gaskin P, Croker SJ, Devos KM, Phillips AL, Hedden P (2006) Function and transcript analysis of gibberellin-biosynthetic enzymes in wheat. *Planta* 223,568-582



- Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815
- Arumuganthan, K, and E.D.Earle (1991) Nuclear DNA content of some important plant species. *Plant Mol. Bio. Rep.* 9, 208-218
- Aylon Y, Kupiec M (2004) New insights into the mechanism of homologous recombination in yeast. *Mutat. Res.* 566, 231-248
- Ayora, S., Piruat, J.I., Luna, R., Reiss, B., Russo, V.E.A., Aguilera, A., and Alonso, J.C. (2002) Characterization of two highly similar Rad51 homologs of *Physcomitrella patens*. *J. Mol. Biol.* 316 35-49
- Azpiroz-Leehan, R. and Feldmann, K.A. (1997) T-DNA insertion mutagenesis in Arabidopsis: Going back and forth. *Trends Genetic* 13, 152-156
- Ball, C.A., Awad, I.A.B., Demeter, J., Gollub, J., Hebert, J.M., Hernandez-Boussard, T., Jin, H., Matese, J.C., Nitzberg, M., Wymore, F., *et al.* (2005) The Stanford Microarray Database accommodates additional microarray platforms and data formats. *Nucleic Acids Res.* 33, 580-582
- Bannister LA, Schimenti JC (2004) Homologous recombinational repair proteins in mouse meiosis. *Cytogenet. Genome. Res.* 107, 191-200
- Barrett, Tugba O. Suzek , Dennis B. Troup , Stephen E. Wilhite , Wing-Chi Ngau , Pierre Ledoux , Dmitry Rudnev , Alex E. Lash , Wataru Fujibuchi and Ron Edgar (2005) NCBI GEO: mining millions of expression profiles - database and tools. *Nucleic Acids Res*, 33, D562-D566
- Bass, H. W., Marshall, W. F., Sedat, J. W., Agard, D. A. and Cande, W. Z. (1997) Telomeres cluster de novo before the initiation of synapsis: a three-dimensional spatial analysis of telomere positions before and during meiotic prophase. *J. Cell Biol.* 137, 5-18
- Bass, H.W., Bordoli, S.J., and Foss, E.M. (2003) The *desynaptic (dy)* and *desynaptic1 (dsy1)* mutations in maize (*Zea mays* L.) cause distinct telomere-misplacement phenotypes during meiotic prophase. *J. Exp. Bot.* 54, 39-46

- Bennett, M.D., Chapman, V. & Riley, R. (1971) The duration of meiosis in pollen mother cells of wheat, rye and triticales. *Proc. Royal Soc., Lond. B*, 178, 259-275
- Bennett, M.D., Finch, R.A., Smith, J.B. and Rao, M.K. (1973) The time and duration of female meiosis in wheat, rye and barley. *Proceeding of the Royal Society of London. Series B. Biological Sciences* 183, 301-319
- Berchowitz LE, Francis KE, Bey AL, Copenhaver GP (2007) The role of AtMUS81 in interference-insensitive crossovers in *A. thaliana*. *PLoS Genetics* 3, 1355–1364
- Betzner, A., Stefan, *et al.* RhoBio. Composition and method for increased meiotic recombination in plants, WO/2002/008432
- Bishop D.K., Zickler D, (2004), Meiotic crossover interference prior to stable strand exchange and synapsis, *Cell* 117, 9-15
- Bishop DK (1994) RecA homologs Dmc1 and Rad51 interact to form multiple nuclear complexes prior to meiotic chromosome synapsis. *Cell* 79, 1081-1092
- Bishop DK, Park D, Xu L, Kleckner N (1992) DMC1: a meiosis-specific yeast homolog of *E.coli* recA required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* 69, 439-456
- Blake NK, Sherman JD, Dvorák J, Talbert LE (2004) Genome-specific primer sets for starch biosynthesis genes in wheat. *Theor Appl Genet* 109, 1295-1302
- Bleuyard, J.Y. and White, C.I. (2004) The Arabidopsis homologue of Xrcc3 plays an essential role in meiosis. *The EMBO Journal* 23, 439-449
- Bleuyard, J.Y., Gallego, M.E. and White, C.I. (2004) Meiotic defects in the Arabidopsis rad50 mutant point to conservation of the MRX complex function in early stages of meiotic recombination. *Chromosoma*, 113, 197–203
- Boden SA, Langridge P, Spangenberg G, Able J (2008) *TaASY1* promotes homologous chromosome interactions and is affected by deletion of *Ph1*. *The Plant Journal* DOI: 10.1111/j.1365-313X.2008.03701.x

- Boden SA, Nadim Shadiac, Elise J Tucker, Peter Langridge and Jason A Able (2007) Expression and functional analysis of *TaASY1* during meiosis of bread wheat (*Triticum aestivum*) *BMC Molecular Biology* 8, 65
- Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R. and Gorlach, J. (2001) Growth stage-based phenotypic analysis of Arabidopsis. *The Plant Cell* 13, 1499-1510
- Brazma A., Hingamp P., Quackenbush J., Sherlock G., Spellman P., Stoeckert C., Aach J., Ansorge W., Ball C. A., Causton H. C., Gaasterland T., Glenisson P., Holstege F. C., Kim I. F., Markowitz V., Matese J. C., Parkinson H., Robinson A., Sarkans U., Schulze-Kremer S., Stewart J., Taylor R., Vilo J., Vingron M (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet.* 29,365-71
- Brazma, A. *et al.*, (2003) ArrayExpress-a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res* 31, 68-71
- Brenneman, M.A., Wagener, B.M., Miller, C.A., Allen, C. and Nickoloff, J.A. (2002) XRCC3 controls the fidelity of homologous recombination: roles for XRCC3 in late stages of recombination. *Mol. Cell* 10, 387-395
- Briggle and Curtis, (1987) L.W. Briggle and B.C. Curtis, Wheat worldwide. In: E.G. Heyne, Editor, *Wheat and Wheat Improvement*, 1-31
- Bundock, P. and Hooykaas, P. (2002) Severe developmental defects, hypersensitivity to DNA-damaging agents, and lengthened telomeres in Arabidopsis MRE11 mutants. *Plant Cell*, 14, 2451-2462
- Bustin, S A, Benes, V, Nolan, T, Pfaffl, M W (2005) Quantitative real-time RT-PCR - a perspective *J Mol Endocrinol* 34, 597-60
- Caldwell G, Nicola McCallum, Paul Shaw, Gary J. Muehlbauer, David F. Marshall and Robbie Waugh (2004) A structured mutant population for forward and reverse genetics in Barley (*Hordeum Vulgare* L.). *The Plant Journal* 40, 143-150

- Callum Mc, C.M., Comai, L., Greene, E.A. & Henikoff, and S. (2000b) Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol.* 123, 439-442
- Callum Mc, C.M., Comai, L., Greene, E.A. & Henikoff, and S. (2000a) Targeted screening for induced mutations. *Nat. Biotechnol.* 18, 455-457
- Cartwright, R, M Tambini, C.E., Simpson, P.J. and Thacker, J, (1998a) The XRCC2 DNA repair gene from human and mouse encodes a novel member of the recA/RAD51 family. *Nucleic Acid Res.* 26, 3084-3089
- Cartwright, R., Dunn, A.M., Simpson, P.J., Tambini, C.E. and Thacker, J, (1998b) Isolation of novel human and mouse genes of the recA/rad51 recombination-repair gene family. *Nucleic Acids Res.* 26, 1653-1659
- Chamakhah, M. and W. Xiao, (1999) Formation of the yeast Mre11-Rad50-Xrs2 complex is correlated with DNA repair and telomere maintenance. *Nucleic Acids Res.* 27, 2072-2079
- Chao S, Sharp PJ, Worland AJ, Koebner RMD, Gale MD (1989) RFLP-based genetic maps of homoeologous group 7 chromosomes. *Theor Appl Genet* 78, 495-504
- Chapman and Riley, (1970) Homoeologous meiotic chromosome pairing in *Triticum aestivum* in which chromosome 5B is replaced by alien homoeologue. *Nature* 226, 376-377
- Chapman, V., Miller, T.E., and Riley, R. (1976) Equivalence of the A genome of bread wheat and that of *Triticum urartu*. *Genet. Res.* 27, 69-76
- Chen Z, Yang H, Pavletich NP (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* 453, 489-484
- Chen ZJ (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annual Review of Plant Biology* 58, 377-406

- Chikashige, Y., Ding, D.-Q., Imai, Y., Yamamoto, M., Haraguchi, T. and Hiraoka, Y. (1997) Meiotic nuclear reorganization: switching the position of centromeres and telomeres in the fission yeast *Schizosaccharomyces pombe*. *EMBO J.* 16, 193-202
- Ciaffi, M. Ciaffi, A.R. Paolacci, E. D'Aloisio, O.A. Tanzarella and E. Porceddu, (2006) Cloning and characterization of wheat PDI (protein disulfide isomerase) homoeologous genes and promoter sequences, *Gene* 366, 209–218
- Clark J, and A. D. Margulies (1965) Isolation and characterization of recombination deficient mutants of *E. coli* K-12. *Proc. Natl Acad. Sci. U.S.A* 53, 61-69
- Close TJ, Wanamaker SI, Caldo RA, Turner SM, Ashlock DA, Dickerson JA, Wing RA, Muehlbauer GJ, Kleinhofs A, Wise RP (2004) A new resource for cereal genomics: 22K barley GeneChip comes of age. *Plant Physiol* 134, 960-968
- Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* 16, 735-743
- Colbert, T., Bradley J. Till, Rachel Tompa, Steve Reynolds, Michael N. Steine, Anthony T. Yeung, Claire M. McCallum, Luca Comai, and Steven Henikoff (2001) High-throughput screening for induced point mutations. *Plant Physiol.* 126, 480-484
- Comai, L. (2000). Genetic and epigenetic interactions in allopolyploid plants. *Plant Mol. Biol.* 43, 387–399
- Copenhaver Gregory P., Kevin C. Keith, and Daphne Preuss (2000) Tetrad analysis in higher plants. A budding technology. *Plant Physiol.* 124, 7–16
- Couteau, F., Belzile, F., Horlow, C., Grandjean, O., Vezon, D. and Doutriaux, M.P. (1999) Random chromosome segregation without meiotic arrest in both male and female meiocytes of a *dmcl* mutant of *Arabidopsis*. *The Plant Cell* 11, 1623-1634
- Cox, T.S., (1998) Deepening the wheat gene pool. *J Crop Prod* 1, 1–25

- Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S (2004) NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. *Nucleic Acids Res* (Database issue) 32, D575–D577
- Crismani W, Baumann U, Sutton T, Shirley N, Webster T, Spangenberg G, Langridge P, Able J. (2006) Microarray expression analysis of meiosis and microsporogenesis in hexaploid bread wheat. *BMC Genomics*. 7
- Cusack BP, Wolfe KH (2007) When gene marriages don't work out: Divorce by subfunctionalization. *Trends Genet* 23, 270–272
- Daboussi, M. J., J. M. Daviere, S. Graziani and T. Langin, (2002) Evolution of the *Fot1* transposons in the genus *Fusarium*: discontinuous distribution and epigenetic inactivation. *Mol. Biol. Evol.* 19, 510–520
- Daphne Preuss Chromatin Silencing and Arabidopsis Development (1999): A Role for Polycomb Proteins. *Plant Cell* 11, 765-768
- Dawe, K. R., Sedat, J. W., Agard, D. A. and Cande, Z. W. (1994) Meiotic chromosome pairing in Maize is associated with a novel chromatin organization. *Cell* 76, 901-912
- Dawe, R.K (1998) Meiotic chromosome organization and segregation in plants. *Annual review of plant physiology & Plant Molecular Biology*. 49, 371-395
- De Bustos, Ruth Pérez and Nicolás Jouve (2007) Characterization of the gene *Mre11* and evidence of silencing after polyploidization in Triticum. *Theoretical and Applied Genetics* 114, 985-999
- Deans, B., Griffin, C.S., O'Regan, P., Jasin, M. and Thacker, J. (2003) Homologous recombination deficiency leads to profound genetic instability in cells derived from *Xrcc2*-knockout mice. *Cancer Res.* 63, 8181–8187

- Delhaize E, Diane M. Hebb, Keith D. Richards, Jian-Ming Lin, Peter R. Ryan, and Richard C. Gardner (1999) Cloning and Expression of a Wheat (*Triticum aestivum* L.) Phosphatidylserine Synthase cDNA. Overexpression in plants alters the compositions of phospholipids *J. Biol. Chem.* 274, 7082-7088
- Deng, Z.Y. and Wang, T. (2007) *OsDMC1* is required for homologous pairing in *Oryza sativa*. *Plant Molecular Biology* 65, 31-42
- Dernburg, A.F., Sedat, J.W., Cande, W.Z. and Bass, H.W (1995) Cytology of telomeres. In *Telomeres*. Blackburn, E.H., and Greider, C.W. (eds.) Cold Spring Harbour, NY: Cold Spring Harbour Laboratories Press, 295-338
- Devos KM, Gale MD (1997) Comparative genetics in the grasses. *Plant Mol Biol* 35, 3-15
- Devos KM, Millan T, Gale MD (1993) Comparative RFLP maps of homoeologous group 2 chromosomes of wheat, rye, and barley. *Theor Appl Genet* 85, 784-792
- DeVries AP (1971) Flowering Biology of Wheat, Particularly in View of Hybrid Seed Production - A Review. *Euphytica* 20,152-170
- Di Capua., E. Di Capua, A. Engel, A. Stasiak and T. Koller, (1982) Characterization of complexes between recA protein and duplex DNA by electron-microscopy, *J. Mol. Biol.* 157, 87-103
- Dickinson HG (1981) The structure and chemistry of plastid differentiation during male meiosis in *Lilium henryi* *J Cell Sci.* 52, 223-41
- Ding, Z.J., Wang, T., Chong, K. and Bai, S. (2001) Isolation and characterization of *OsDMC1*, the rice homologue of the yeast *DMC1* gene essential for meiosis. *Sex Plant Reproduction* 13, 285-288
- Dirks, R.H.G. *et al.* Rijk Zwaan Zaadteelt en Zaadhandel B.V. Reverse breeding, WO/2003/017753
- Dong CM, Whitford R, Langridge P (2002) A DNA mismatch repair gene links to the *Ph2* locus in wheat. *Genome* 45, 116-124



- Dosanjh, M.K., Collins, D.W. , Fan, W., Lennon, G.G., Albala, J. S., Shen, Z and Schild, D.(1998) Isolation and characterization of RAD51C, a new human member of the RAD51 family of related genes. *Nucleic Acids Res.* 26, 1179-1184
- Doutriaux, M.P., Couteau, F., Bergounioux, C. and White, C. (1998) Isolation and characterisation of the *RAD51* and *DMC1* homologs from *Arabidopsis thaliana*. *Molecular and General Genetics* 257, 283-291
- Dray Eloïse, Nicolas Siaud, Emeline Dubois, and Marie-Pascale Doutriaux (2006) Interaction between Arabidopsis Brca2 and Its Partners Rad51, Dmc1, and Dss1. *Plant Physiol.* 140, 1059-1069
- Driscoll, C.J. (1972) Genetic suppression of homoeologous chromosome pairing in hexaploid wheat. *Can. J. Genet. Cytol.* 14, 39-42
- Duarte, M., Liying Cui, P. Kerr Wall, Qing Zhang, Xiaohong Zhang, Jim Leebens-Mack, Hong Ma, Naomi Altman, and Claude W. dePamphilis (2006) Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of Arabidopsis. *Mol Biol Evol* 23, 469–478
- Dubcovsky J, Ramakrishna W, San Miguel PJ, Busso CS, Yan L, Shiloff BA, Bennetzen JL (2001) Comparative sequence analysis of collinear barley and rice bacterial artificial chromosomes. *Plant Physiol* 125, 1342-1353
- Dudas, A. and Chovanec, M. (2004) DNA double-strand break repair by homologous recombination. *Mutant Research* 566, 630-670
- Durrant WE, Dong X (2004) *Systemic acquired resistance*. *Annu Rev Phytopathol* 42,185–209
- Dvořák, J., Luo, M.C., Yang, Z.L., and Zhang, H.B. (1998) The structure of the *Aegilopes tauschii* genepool and the evolution of hexaploid wheat. *Theor. Appl. Genet.* 97, 657–670
- Edgar R, Domrachev M, Lash AE. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210

- Ehrendorfer, F.L. (1980) Polyploidy and distribution. In *Polyploidy-Biological Relevance*, W.H. Lewis, ed (New York: Plenum Press), 45–60
- Endo TR, Gill BS (1996) The deletion stocks of common wheat. *J Heredity* 87, 295–307
- Erik Wijnker and Hans de Jong (2008) Managing meiotic recombination in plant breeding. *Trends in plant sciences* 16, 640-646
- Evans, L.T., Bingham, J. & Roskams, M.A. (1972) The pattern of grain set within ears of wheat. *Austr. J. Biol. Sci.*, 25, 1-8
- Farmer, J.B. and Moore, J.E.S. (1905) On the meiotic phase (reduction divisions) in animals and plants. *Journal of Microscopical Science* 48, 489-557
- Feldman, M. (1966) The effect of chromosomes 5<sup>B</sup>, 5<sup>D</sup>, and 5 on chromosomal pairing in *Triticum aestivum*. *Proc Nat Acad Sci* 55, 1147-1453
- Feldman, M., B. Liu, G. Segal, S. Abbo, A.A. Levy, and J.M. Vega. (1997) Rapid elimination of low-copy DNA sequences in polyploid wheat: A possible mechanism for differentiation of homoeologous chromosomes. *Genetics* 147, 1381–1387
- Feldmann, K.A. (1991) T-DNA insertion mutagenesis in *Arabidopsis* mutational spectrum. *The Plant Journal* 1, 71-82
- Filippo, J.S., Sung, P. and Klein, H. (2008) Mechanism of eukaryotic homologous recombination. *Annual Review of Biochemistry* 77, 229-257
- Force, Allan, Lynch, Michael, Pickett, F. Bryan, Amores, Angel, Yan, Yi-lin, Postlethwait, John (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531–1545
- Francis K.E.*et al.*, (2006) Separation of *Arabidopsis* pollen tetrads is regulated by *QUARTET1*, a pectin methylesterase gene, *Plant Physiol.* 142, 1004–1013
- Franklin, A.E., Golubovskaya, I.N., Bass, H.W., and Cande, W.Z. (2003) Improper chromosome synapsis is associated with elongated RAD51 structures in maize *desynaptic2* mutant. *Chromosoma* 10.1007/S00412-003-0242-8

- Franklin, A.E., McElver, J., Sunjevaric, I., Rothstein, R., Bowen, B. and Cande, W.Z. (1999) Three-dimension microscopy of the *RAD51* recombination protein during meiotic prophase. *The Plant Cell* 11, 809-824
- Fu *et al.*, D. Fu, C. Uauy, A. Blechl and J. Dubcovsky, (2007) RNA interference for functional gene analysis, *Transgenic Research* 16 ,689–701
- Fujita, N. and Taira, T., (1998) A 56 kDa protein is a novel granule-bound starch synthase existing in the pericarps, aleurone layers, and embryos of immature seed in diploid wheat (*Triticum monococcum* L.). *Planta* 207,125–132
- Fuller LF, Painter RB (1988) A Chinese hamster ovary cell line hypersensitive to ionizing radiation and deficient in repair replication. *Mutat Res.* 193,109–121
- Galbraith, D.W., and Birnbaum, K. (2006) Global studies of cell type-specific gene expression in plants. *Annual Reviews of Plant Biology* 57, 451-475
- Gerton JL, Hawley RS. (2005) Homologous chromosome interactions in meiosis: diversity amidst conservation. *Nat Rev Genet.* 6, 477-487
- Gill, K. S. and B. S. Gill, (1996) A PCR-based screening assay of *Ph1*, the chromosome pairing regulator gene of wheat. *Crop Sci.* 36,719-722
- Gill, K. S., B. S. Gill, T. R. Endo, and Y. Mukai, (1993) Fine physical mapping of *Ph1*, a chromosome pairing regulator gene in polyploid wheat. *Genetics* 134, 1231-1235
- Gollub, J., Ball,C.A., Binkley,G., Demeter,J., Finkelstein,D.B., Hebert,J.M., Hernandez-Boussard,T., Jin,H., Kaloper,M., Matese, J.C., Michael Nitzberg, Farrell Wymore, Zachariah K. Zachariah, Patrick O. Brown, and Gavin Sherlock (2003) The Stanford Microarray Database: data access and quality assessment tools. *Nucleic Acids Res.* 31, 94–96
- Golubovskaya, I.N., Grebennikova, Z.K., Auger, D.L., and Sheridan, W.F. (1997) The maize desynaptic1 mutation disrupts meiotic chromosome synapsis. *Dev. Genet.* 21, 146–159

- Grelon M, Vezon D, Gendrot G, Pelletier G (2001) *AtSPO11-1* is necessary for efficient meiotic recombination in plants. *EMBO J.* 20, 589-600
- Grelon, M., Gendrot, G., Vezon, D., Pelletier, G., Mathilde, G., Ghislaine, G., Daniel, V. and Georges, P. (2003). The Arabidopsis MEI1 gene encodes a protein with five BRCT domains that is involved in meiosis-specific DNA repair events independent of SPO11-induced DSBs. *Plant J.* 35, 465–475
- Griffiths S, Sharp R, Foote TN, Bertin I, Wanous M, Reader S, Colas I, Moore G (2006) Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature* 439,749–752
- Gupta P. Balyan H, Edwards K, Isaac P. Korzun V, Roder M, Gautier M-F, Joudrier P, Schlatter A, Dubcovsky J, De la Pena R, Khairallah M, Penner G, Hayden M, Sharp P, Keller B, Wang R, Hardouin J, Jack P, Leroy P (2002) Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat. *Theor Appl Genet* 105, 413-422
- Gupta, R.C., Golub, E., Bi, B., Radding, C.M. (2001) The synaptic activity of HsDmc1, a human recombination protein specific to meiosis. *Proc. Natl. Acad. Sci.* 98, 8433–8439
- Habu T, Taki T, West A, Nishimune Y, Morita T. (1996) The mouse and human homologs of *DMC1*, the yeast meiosis-specific homologous recombination gene, have a common unique form of exon-skipped transcript in meiosis. *Nucleic Acids Research* 24, 470–477
- Haga H, Yamada R, Ohnishi Y, Nakamura Y, Tanaka T (2002) Gene-based SNP discovery as part of the Japanese Millenium Genome Project: identification of 190,562 genetic variations in the human genome. *J Hum Genet* 47, 605-610
- Hanin, M., Volrath, S., Bogucki, A., Briker, M., Ward, E., and Paszkowski, J. (2001) Gene targeting in Arabidopsis. *Plant J.* 28, 671–677
- Hartung, F. and Puchta, H. (2001) Molecular chracterisation of homologues of both subunits A (*SPO11*) and B of the archaebacterial topoisomerase 6 in plants. *Gene* 271, 81-86

- He, P., B. R. Frebe, B. S. Gill and J. M. Zhou, (2003) Allopolyploidy alters gene expression in the highly stable hexaploid wheat. *Plant Mol. Biol.* 52, 401–414
- Hegarty, M. J., J. M. Jones, I. D. Wilson, G. L. Barker, J. A. Coghill *et al.*, (2005) Development of anonymous cDNA microarrays to study changes to the Senecio floral transcriptome during hybrid speciation. *Mol. Ecol.* 14, 2493–2510
- Henikoff, S., Till, B.J. & Cornai, L. (2004) TILLING. Traditional mutagenesis meets functional genomics. *Plant Physiol.* 135, 630-636
- Higgins, J. D., E. Sanchez-Moran, S. J. Armstrong, G. H. Jones and F. C. H. Franklin, (2005) The Arabidopsis synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. *Genes Dev.* 19, 2488–2500
- Hilu, K.W. (1993) Polyploidy and the Evolution of Domesticated Plants *American Journal of Botany.* 80, 1494-1499
- Höfgen R, Axelsen KB, Kannangara CG, Schuttke I, Pohlenz HD, Willmitzer L, Grimm B, Wettstein DV (1994) A visible marker for antisense mRNA expression in plants: inhibition of chlorophyll synthesis with a glutamate-1-semialdehyde aminotransferase antisense gene. *Proc Natl Acad Sci USA* 91, 1726–1730
- Holliday .R and R. E. Halliwell, (1968) An endonuclease deficient strain of *Ustilago maydis* *Genetic Res.* 12, 95-98
- Holliday, R.A. (1964) A mechanism for gene conversion in fungi. *Genetics Research Cambridge* 5, 282-304
- Hollingsworth, N. M., and B. Byers (1989) *HOP1*: a yeast meiotic pairing gene. *Genetics* 121, 445–462
- Hollingsworth NM, Byers B (1990) The HOP1 gene encodes a meiosis-specific component of yeast chromosomes. *Cell.* 61, 75-84.

- Hsaio C, Chatterton NJ, Asay KH, Jensen KB (1994) Phylogenetic relationships of 10 grass species: an assessment of phylogenetic utility of the internal transcribed spacer region in nuclear ribosomal DNA in monocots. *Genome* 37, 112–120
- Huang, S., A. Sirikhachornkit, X. Su, J. Faris, *et al.*, (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat. *Proceedings of the National Academy of Sciences* 99, 8133–8138
- Ivancic-Bace I, Peharec P, Moslavac S, Skrobot N, Salaj-Smic E, Brcic-Kostic K (2003) RecFOR function is required for DNA repair and recombination in a recA loading-deficient recB mutant of *Escherichia coli*. *Genetics* 163, 485–494
- J.D. Higgins, S. J. Armstrong, F. C. H. Franklin and G. H. Jones (2004) The *Arabidopsis* MutS homolog AtMSH4 functions at an early step in recombination: evidence for two classes of recombination in *Arabidopsis*, *Genes Dev.* 18, 2557–2580
- Jain SK (1975) Population structure and the effects of breeding system. In: Frankel, O.H. & Hawkes, J.G. (eds.) Crop Genetic Resources for Today and Tomorrow. *Cambridge Univ. Press.* 15–36
- Ji, L.H. and Langridge, P. (1994) An early meiosis cDNA clone from wheat. *Mol. Gen. Genet.* 243, 17–23
- Jones NJ, Cox R, Thacker J (1987) Isolation and cross-sensitivity of X-ray-sensitive mutants of V79-4 hamster cells. *Mutat Res.* 183, 279–286
- Kashkush, K., M. Feldman and A. A. Levy, (2003) Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* 33, 102–106
- Kashukush, K., M. Feldman and A. A. Levy, (2002) Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics* 160, 1651–1659

- Kathiresan, A., Khush, G.S. and Bennett, J. (2002) Two rice *DMC1* genes are differentially expressed during meiosis and during haploid and diploid mitosis. *Sex Plant Reproduction* 14, 257-267
- Keeney S.(2001) Mechanism and control of meiotic recombination initiation. *Curr. Top. Dev. Biol.* 52, 1-53
- Kellogg EA (1998) Relationships of cereal crops and other grasses. *Proc Natl Acad Sci USA* 95, 2005–2010
- Kellogg EA (2001) Evolutionary history of the grasses. *Plant Physiol* 125, 1198–1205
- Kempin, S.A., Liljegren, S.J., Block, L.M., Rounsley, S.D., Lam, E. and Yanofsky, M.F. (1997) Inactivation of the Arabidopsis AGL5 MADS-box gene by homologous recombination. *Nature* 389, 802-803
- Kerzendorfer, C., Vignard, J., Pedrosa-Harand, A., Siwiec, T., Akimcheva, S., Jolivet, S., Sablowski, R., Armstrong, S., Schweizer, D., Mercier, R., and Schlogelhofer, P. (2006) The *Arabidopsis thaliana* MND1 homologue plays a key role in meiotic homologous pairing, synapsis and recombination. *J.Cell Sci.* 119 2486–2496
- Khoo, K.H.P, Jolly, H.R. and Able, J.A. (2008) The *RAD51* gene family in bread wheat is highly conserved across eukaryotes, with *RAD51A1* unregulated during early meiosis. *Functional Plant Biology* 35, 1267-1277
- Kihara (1924) Cytologische und genetische Studien bei wichtigen Getreidearten mit besondere Riichsicht auf das Verhalten der Chromosomen und die Sterilitat in den Bastarden. *Mem. Coll. Sci.* Kyoto Univ. 200 pp.
- Kihara, H., (1944). Discovery of the DD-analyzer, one of the ancestors of *vulgare* wheats. *Nogyo oyobi Engei (Agric.&Hortic.)* 19, 889–890 (in Japanese)
- Kirby, E.J.M. & Appleyard, M. (1987) *Cereal development guide*. Stoneleigh, Kenilworth, UK, NAC Cereal Unit. 85pp



- Klimyuk, V.I and Jones, J.D. (1997) *AtDMC1*, the Arabidopsis homologue of the yeast DMC1 gene: characterization, transposon-induced allelic variation and meiosis- associated expression. *Plant J.* 11, 1-14
- Klutstein, Michael, Shaked, Hezi, Sherman, Amir, Avivi-Ragolsky, Naomi, Shema, Efrat, Zenvirth, Drora, Levy, Avraham A., Simchen, Giora (2008) Functional Conservation of the Yeast and Arabidopsis RAD54-Like Genes. *Genetics* 178, 2389-2397
- Koayashi, T., Kobayashi, E., Sato, S., Hotta, Y., Miyajima, N., Tanaka, A. and Tabata, S. (1994) Characterization of cDNAs induced in meiotic prophase in lily microporocytes. *DNA Res.* 1, 15-26
- Kobayashi, T., Hotta, Y., and Tabata, S. (1993) Isolation and characterization of a yeast gene that is homologous with a meiosis specific cDNA from a plant. *Mol. Gen. Genet.* 237, 225–232
- Koller, B.H., Hagemann, L.J., Doetschman, T., Hagaman, J.R., Huang, S., Williams, P.J., First, N.L., Maeda, N. and Smithies, O. (1989) Germ-line transmission of a planned alteration made in a hypoxanthine phosphoribosyltransferase gene by homologous recombination in embryonic stem cells. *PROC. NATL. ACAD. SCI. USA* 86, 8927-8931
- Kolodner R. (1996) Biochemistry and genetics of eukaryotic mismatch repair. *Genes & Dev.* 10, 1433–1442
- Koornneef, M., Dellaert, L.W.M. and VAN DER Veen, J.H. (1982) EMS- and radiation – induced mutation frequencies at individual loci in *Arabidopsis thaliana* (L.) Heynh. *Mutat. Res.* 93, 109-123
- Korber-Grohne, U (1988) Microscopic methods of identification of plant fibres and animal hairs from the prince's tomb of Hochdorf, Southern Germany. *Journal of Archaeological science* 15, 73-82
- Krysan, P.J., Young, J.C., Tax, F., and Sussman, M.R. (1996) Identification of transferred DNA insertions within *Arabidopsis* genes involved in signal transduction and ion transport. *Proc. Natl. Acad. Sci. USA* 93, 8145-8150

- Kurata N, Moore G, Nagamura Y, Foote T, Yano M, Minobe Y, Gale M. (1994) Conservation of genome structure between rice and wheat. *Biotechnology* 12, 276–278
- Kurber-Grohne, U. (1988). Nutzpflanzen in Deutschland. Theiss, Stuttgart
- Kurumizaka, H., Enomoto, R., Nakada, M., Eda, K., Yokoyama, S. and Shibata, T. (2003) Region and amino acid residues required for Rad51C binding in the human Xrcc3 protein. *Nucleic Acids Res.* 31, 4041–4050
- Kurumizaka, H., Ikawa, S., Nakada, M., Eda, K., Kagawa, W., Takata, M., Takeda, S., Yokoyama, S. and Shibata, T. (2001) Homologous-pairing activity of the human DNA-repair proteins Xrcc3.Rad51C. *Proc. Natl Acad. Sci. USA*, 98, 5538–5543
- Kurumizaka, H., Ikawa, S., Nakada, M., Enomoto, R., Kagawa, W., Kinebuchi, T., Yamazoe, M., Yokoyama, S. and Shibata, T. (2002) Homologous pairing and ring and filament structure formation activities of the human Xrcc2\*Rad51D complex. *J. Biol. Chem.* 277, 14315–14320
- Lai, Z., B.L. Gross, Y. Zou, J. Andrews, and L.H. Rieseberg (2006) Microarray analysis reveals differential gene expression in hybrid sunflower species. *Mol. Ecol.* 15, 1213–1227
- Lancashire, P.D., H. Bleiholder, P. Langelüddeke, R. Strauss, T. Van den Boom, E. Weber, and A. Witzemberger (1991) A uniform decimal code for growth stages of crops and weeds. *Ann. Appl. Biol.* 119,561–601
- Leader D.J. (2005) Transcriptional analysis and functional genomics in wheat. *Journal of Cereal Science* 41, 149-163
- Lees-Miller, S.P. and Meek, K. (2003) Repair of DNA double strand breaks by non-homologous end joining. *Biochimie.* 85, 1161-1173.
- Leitch IJ, Bennett MD (1997) Polyploidy in angiosperms. *Trends in Plant Science* 2, 470-476
- Leitch IJ, Soltis DE, Soltis PS, Bennett MD (2005) Evolution of DNA amounts across land plants (Embryophyta). *Annals of Botany* 95, 207–217.
- Levin, D.A. (1983) Polyploidy and novelty in flowering plants. *Am. Nat.* 122, 1–25

- Li Wuxing, Xiaohui Yang, Zhenguo Lin, Ljudmilla Timofejeva, Rong Xiao, Christopher A. Makaroff, and Hong Ma (2005) The *AtRAD51C* Gene Is Required for Normal Meiotic Chromosome Synapsis and Double-Stranded Break Repair in Arabidopsis. *Plant physiol* 138, 965-976
- Li X, Song Y, Century K, Straight S, Ronald P, Dong X, Lassner M, Zhang Y (2001) A fast neutron deletion mutagenesis-based reverse genetics system for plants. *Plant J* 27,235–242
- Li, J., Harper, L.C., Golubovskaya, I., Wang, C.R., Weber, D., Meeley, R.B., McElve, J., Bowen, B., Cande, W.Z. and Schnable, P.S. (2007) Functional analysis of maize *RAD51* in meiosis and double-strand break repair. *Genetics* 176, 1469-1482
- Li, W., Changbin, C., Ullrich, M.M., Ljudmilla, T., Elmon, S., Ma, H. and Reiss, R. (2004) The Arabidopsis *AtRAD51* gene is dispensable for vegetative development but required for meiosis. *PROC. NATL. ACAD. SCI. USA* 101, 10596-10601
- Lichten M (2001) Meiotic recombination: breaking the genome to save it. *Curr. Biol.* 11, 253-56
- Lin, Z., Kong, H., Nei, M. and Ma, H. (2006) Origins and evolution of the *recA/RAD51* gene family: Evidence for ancient gene duplication and endosymbiotic gene transfer. *PROC. NATL. ACAD. SCI. USA* 103, 10328-10333
- Lio, Y.C., Mazin, A.V., Kowalczykowski, S.C. and Chen, D.J. (2003) Complex formation by the human Rad51B and Rad51C DNA repair proteins and their activities in vitro. *J. Biol. Chem.* 278, 2469–2478
- Lipshutz R. J., Fodor S. P., Gingeras T. R., Lockhart D. J. (1999) High density synthetic oligonucleotide arrays. *Nat Genet.* 21, 20–4
- Lishuang Shen, Jian Gong, Rico A. Caldo, Dan Nettleton, Dianne Cook, Roger P. Wise, and Julie A. Dickerson (2005) BarleyBase-an expression profiling database for plant genomics *Nucl. Acids Res.* 33, D614-D618
- Liu and West, (2004) Happy Hollidays: 40th anniversary of the Holliday junction. *Nature Reviews Molecular Cell Biology* 5, 937-944

- Liu, B., J.M. Vega, G. Segal, S. Abbo, M. Rodova, and M. Feldman. (1998) Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. I. Changes in low-copy noncoding DNA sequences. *Genome* 41, 272–277
- Liu, N, M Lamerdin, J.E., Tebbs, R.S. *et al.* (1998) XRCC2 and XRCC3, new human Rad51-family members, promote chromosome stability and protect against DNA cross-links and other damages. *Mol Cell.* 1, 783-793
- Liu, W.M., Mei, R., Di, X., Ryder, T.B., Hubbell, E., Dee, S., Webster, T.A., Harrington, C.A., Ho, M.H., Baid, J. and Smeekens, S.P. (2002) Analysis of high density expression microarrays with signed-rank call algorithms. *Bioinformatics* 18, 1593-1599
- Liu, Y., Masson, J.Y., Shah, R., O'Regan, P. and West, S.C (2004) RAD51C is required for Holliday junction processing in mammalian cells. *Science* 303, 243–246
- Liu, Z. and Adams, K. L. (2007) Expression partitioning between genes duplicated by polyploidy under abiotic stress and during organ development. *Current Biology* 17, 1669-1674
- Loidl, J, Klein, F and Scherton,H. (1994) Homologous pairing is reduced but not abolished in asynaptic mutant yeast. *J.Cell.Bio.* 125, 1195-1200
- Loukoianov A, Yan L, Blechl A, Sanchez A, Dubcovsky J. (2005) Regulation of *VRN-1* vernalization genes in normal and transgenic polyploid wheat. *Plant Physiol.*138, 2364–2373
- Lovett, S. T., and R. K. Mortimer, (1987) Characterization of null mutants of the *RAD55* gene of *Saccharomyces cerevisiae*: effects of temperature, osmotic strength and mating type. *Genetics* 116, 547–553
- Lusetti and Michael M. Cox (2002) The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annual Review of Biochemistry* 71, 71-100
- Lynch and Conery, (2000) The evolutionary fate and consequences of duplicate genes, *Science* 290, 1151–1155

- Lynch, M. and A. Force (2000) The evolutionary fate and consequences of duplicate genes. *Science* 290, 1151-1155
- Ma H., (2006) Molecular portrait of *Arabidopsis* meiosis. In: C.R. Somerville and E.M. Meyerowitz, Editors. *The Arabidopsis Book*, American Society of Plant Biologists, 1-39
- Madlung, A., R. W. Masuelli, B. Watson, S. H. Reynolds, J. Davidson *et al.*, (2002) Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis* allotetraploids. *Plant Physiol.* 129, 733-746
- Markmann-Mulisch, U., Hadi, M.Z., Koepchen, K., Alonso, J.C., Russo, V.E.A., Schell, J., and Reiss, B. (2002) The organization of *Physcomitrella patens* RAD51 genes is unique among eukaryotic organisms. *Proc. Natl. Acad. Sci. USA* 99 2959-2964
- Martienssen, R.A (1998). Functional genomics: probing plant gene function and expression with transposons. *Proceedings of the National Academy of Science of the United States of America* 95, 2021-2026
- Masson J.E. and J. Paszkowski, (1997) *Arabidopsis thaliana* mutants altered in homologous recombination, *Proc. Natl. Acad. Sci. U. S. A.* 94, 11731-11735
- Masson J.Y., Davies A.A., Hajibagheri N., Van Dyck E., Benson F.E., Stasiak A.Z., Stasiak A., West S.C. (1999) The meiosis-specific recombinase hDmc1 forms ring structures and interacts with hRad51. *EMBO J.* 18:6552-6560
- Masson, J. E., King, P. J., Paszkowski, J. (1997) Mutants of *Arabidopsis thaliana* Hypersensitive to DNA-Damaging Treatments *Genetics* 146, 401-407
- Masson, J.-Y., and West, S. C. (2001). The RAD51 and DMC1 recombinases: A non-identical twin relationship. *Trends Biochem. Sci.* 26, 131-136

- Masson, J.Y., Tarsounas, M.C., Stasiak, A.Z., Stasiak, A., Shah, R., McIlwraith, M.J., Benson, F.E. and West, S.C. (2001) Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev.* 15, 3296–3307
- Masterson, J (1994) Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* 264, 421–424
- Matzke *et al.*, (1994) "Homology-dependent gene silencing in transgenic plants: epistatic silencing loci contain multiple copies of methylated transgenes," *Mol. Gen. Genet* 244, 219–229
- Mazin *et al.*, A.V. Mazin, C.J. Bornarth, J.A. Solinger, W.D. Heyer and S.C. Kowalczykowski, (2000) Rad54 protein is targeted to pairing loci by the Rad51 nucleoprotein filament, *Mol. Cell* 6, 583–592
- McFadden, E.S. & E.R. Sears, (1946) The origin of *Triticum spelta* and its free-threshing hexaploid relatives. *J Hered* 37, 81–90, 107–116
- McKinney, E.C., Ali, N., Traut, A., Feldmann, K.A., Belostotsky, D.A., McDowell, J.M., and Meagher, R.B. (1995) Sequence-based identification of T-DNA insertion mutations in *Arabidopsis*: Actin mutants act2-1 and act4-1. *Plant J.* 8, 613–622
- McLeish, J. and Snoad, B. (1958) *Looking at Chromosomes*: St Martin's, Macmillan.
- Mello-Sampayo, T. and Lorente, R. (1968) The role of chromosome 3D in the regulation of meiotic pairing in hexaploid wheat. *EWAC Newslett.* 2, 16–24
- Mello-Sampayo, T. (1971) Genetic regulation of meiotic chromosomes pairing by chromosome 3D of *Triticum aestivum*. *Nat. New Biol.* 230, 23–24
- Mello-Sampayo, T. and Canas, P, (1973) Suppressors of meiotic chromosome pairing in common wheat. In *Proceedings of the Fourth International Wheat Genetics Symposium*. Columbia, Missouri, pp. 709–713
- Mercier R, and Grelon M, (2008) Meiosis in plants: ten years of gene discovery, *Cytogenet. Genome Res.* 120, 281–290

- Meselson, M.S. and Radding, C.M. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 358-361
- Mikhailova EI, Phillips D, Sosnikhina SP, Lovtsyus AV, Jones RN, Jenkins G. (2006) Molecular assembly of meiotic proteins Asy1 and Zyp1 and pairing promiscuity in rye (*Secale cereale* L.) and its synaptic mutant sy10. *Genetics* 174, 1247–1258
- Miller KA, Sawicka D, Barsky D, Albala JS (2004) Domain mapping of the RAD51 paralog protein complexes. *Nucleic Acids Research* 32, 169–178
- Miller, K.A., Yoshikawa, D.M., McConnell, I.R., Clark, R., Schild, D. and Albala, J.S. (2002) RAD51C interacts with RAD51B and is central to a larger protein complex in vivo exclusive of RAD51. *J. Biol. Chem.* 277, 8406–8411
- Mochida K, Yamazaki Y, Ogihara Y (2004). Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Mol Genet Genomics* 270, 371–377
- Mochida, K., Y. Yamazaki, and Y. Ogihara (2003) Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Molecular Genetic Genomics* 270, 371-377
- Modrich, P., and Lahue, R. S. (1996) Mismatch repair in replication fidelity, genetic recombination and cancer biology *Annu. Rev. Biochemistry* 65, 101-133
- Morimatsu K, Kowalczykowski SC (2003) RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Mol Cell* 11, 1337–1347
- Morimoto R, Kosugi T, Nakamura C, Takumi S (2005) Intragenic diversity and functional conservation of the three homoeologous loci of the KN1-type homeobox gene Wknx1 in common wheat. *Plant Mol Biol* 57,907–924



- Mujeeb-Kazi, A., V. Rosas, and S. Roldan. (1996) Conservation of the genetic variation of *Triticum tauschii* (Coss.) Schmalh. (*Aegilops squarrosa* auct. non L.) in synthetic hexaploid wheats (*T. Turgidum* L. s.lat. X *T. tauschii*;  $2n = 6x = 42$ , AABBDD) and its potential utilization for wheat improvement. *Genetic Resources and Crop Evolution* 43, 129-134
- Mutti JS, Gill KS (The 11th International Wheat Genetics Symposium proceedings, 2008)
- Mysore *et al* (2000). "An Arabidopsis histone H2A mutant is deficient in Agrobacterium T-DNA integration", *PROC. NATL. ACAD. SCI. USA* 97, 948-953
- Naito K, Kusaba M, Shikazono N, Takano T, Tanaka A, Tanisaka T, Nishimura M (2005) Transmissible and nontransmissible mutations induced by irradiating *Arabidopsis thaliana* pollen with gamma -rays and carbon ions. *Genetics* 169, 881–889
- Nara, T., Yamamoto, T. and Sakaguchi, K. (2000) Characterization of interaction of C- and N-terminal domains in *LIM/DMC1* and *RAD51* from basidiomycetes, *Coprinus cinereus*. *Biochemical and Biophysical Research Communications* 275, 97-102
- New J.H., Sugiyama, T., Zaitseva, E. and Kowalczykowski, S.C. (1998) Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature*, 391, 407–410
- Nomura T, Ishihara A, Yanagita RC, Endo TR, Iwamura H (2005) Three genomes differentially contribute to the biosynthesis of benzoxazinones in hexaploid wheat. *Proc Natl Acad Sci USA* 102, 16490–16495
- Novak, S.J., Soltis, D.E., and Soltis, P.S. (1991) Ownbey's *Tragopogons*: 40 years later. *Am. J. Bot.* 78, 1586–1600
- Ogawa T, Wabiko H, Tsurimoto T, Horii T, Masukata H, Ogawa H. (1979) Characteristics of purified recA protein and the regulation of its synthesis in vivo. *Cold Spring Harb. Symp. quant. Biol.* 43, 909–915

- Ogawa, H., K. Johzuka, T. Nakagawa, S. H. Leem, and A. H. Hagihara, (1995) Functions of the yeast meiotic recombination genes, *MRE11* and *MRE2*. *Adv. Biophys.* 31, 67-76
- Ohno S (1970) Evolution by gene duplication. eds Allen George, London Unwin(Springer-Verlag, New York)
- Orr-Weaver, T.L(1995) Meiosis in *Drosophila*: Seeing is believing. *Proc. Natl. Acad. Sci. USA.* 92, 1.443-1.449
- Osakabe, K., Yoshioka, T., Ichikawa, H. and Toki, S. (2002) Molecular cloning and characterization of *RAD51*-like genes from *Arabidopsis thaliana*. *Plant Molecular Biology* 50, 69-79
- Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, Lee HS, *et al.* (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics* 19, 141–147
- Osman, K., E. Sanchez-Moran, J. D. Higgins, G. H. Jones and F. C. H. Franklin, (2006) Chromosome synapsis in *Arabidopsis*: analysis of the transverse filament protein ZYP1 reveals novel functions for the synaptonemal complex. *Chromosoma* 115, 212–219
- Ozkan, H., Levy, A.A., and Feldman, M. (2001) Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops*–*Triticum*) group. *Plant Cell* 13, 1735–1747
- Paques, Frederic, Haber, James E (1999) Multiple Pathways of Recombination Induced by Double-Strand Breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev* 63, 349-404
- Passy S.I., Yu X., Li Z., Radding C.M., Masson J.Y., West S.C., Egelman E.H.(1999) *Human Dmc1 protein binds DNA as an octameric ring*. *Proc. Natl. Acad. Sci. U.S.A* 96, 10684–10688
- Pawlowski, W.P. and Cande, W.Z. (2005) Coordinating the events of the meiotic prophase. *Trends in Cell Biology* 15, 674-681
- Peach, C. & Velten, J.(1991) Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. *Plant Molecular Biology* 17, 49-60

- Pradillo M .et al., (2007) An analysis of univalent segregation in meiotic mutants of *Arabidopsis thaliana*: a possible role for synaptonemal complex. *Genetics* 175, 505–511
- Pestova E, Ganai MW, Roder MS (2000) Isolation and mapping of microsatellite markers specific for the D genome of bread wheat. *Genome* 43, 689-697
- Peterson, G and O.SEBERG (2005) Phylogenetic Analysis of Allopolyploid Species *Czech J. Genet. Plant Breed* 41
- Peterson, G., O. Seberg, M. Yde, & K. Berthelsen (2006) Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). *Molecular Phylogenetics and Evolution* 39, 70-82
- Petukhova, G. Petukhova, S. Stratton and P. Sung (1998) Catalysis of homologous DNA pairing by yeast Rad51 and Rad54 proteins. *Nature* 393, 91–94
- Petukhova, G. Petukhova, S. Van Komen, S. Vergano, H. Klein and P. Sung (1999) Yeast Rad54 promotes Rad51-dependent homologous DNA pairing via ATP hydrolysis-driven change in DNA double helix conformation, *J. Biol. Chem.* 274, 29453–29462
- Pittman DL, Schimenti JC (2000) Midgestation lethality in mice deficient for the RecA-related gene, Rad51d/Rad51l3. *Genesis* 26, 167–173
- Pittman, D.L., Cobb, J., Schimenti, K.J., Wilson, L.A., Cooper, D.M., Brignull, E., Handel, M.A., and Schimenti, J.C. (1998) Meiotic prophase arrest with failure of chromosome synapsis in mice deficient for Dmcl, a germline-specific RecA homolog. *Mol. Cell* 1, 697-705
- Pittman, D.L., Weinberg, L.R., and Schimenti, J.C. (1998) Identification, characterization, and genetic mapping of Rad51d, a new mouse and human RAD51/recA- related gene. *Genomics.* 49, 103-111
- Poole R, Barker G, Wilson ID, Coghill JA, Edwards KJ (2007). Measuring global gene expression in polyploidy; a cautionary note from allohexaploid wheat. *Funct Integr Genomics*, 207–219

- Prinz S, Klein F, Auer H, Schweizer D, Primig M. A (1995) DNA binding factor (UBF) interacts with a positive regulatory element in the promoters of genes expressed during meiosis and vegetative growth in yeast. *Nucleic Acids Res.* 23, 3449–3456
- Puchta, H., Dujon, B., and Hohn, B. (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc. Natl. Acad. Sci. USA* 93, 5055–5060
- Puizina, J., Siroky, J., Mokros, P., Schweizer, D. and Riha, K. (2004) Mre11 deficiency in Arabidopsis is associated with chromosomal instability in somatic cells and Spo11-dependent genome fragmentation during meiosis. *Plant Cell*, 16, 1968–1978
- Ramakrishna W, Dubcovsky J, Park YJ, Busso C, Embereton J, SanMiguel P, Bennetzen JL (2002) Different types and rates of genome evolution detected by comparative sequence analysis of orthologue segments from four cereal genomes. *Genetics* 162, 1389-1400
- Ramsey, J., and D.W. Schemske (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annu. Rev. Ecol. Syst.* 29, 467–501
- Rastogi S, Liberles DA (2005) Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol Biol* 5, 28
- Rensink W.A and Robin Buel C, (2005) Arabidopsis to Rice. Applying Knowledge from a Weed to Enhance Our Understanding of a Crop Species. *Plant physiol* 135, 622-629
- Rhee SY, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A, Garcia-Hernandez M, Huala E, Lander G, Montoya M, Miller N, Mueller LA, Mundodi S, Reiser L, Tacklind J, Weems DC, Wu Y, Xu I, Yoo D, Yoon J, Zhang P (2003) The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic Acids Research* 31, 224

- Rice, M.C, M Smith, S.T., Bullrich, F, M Havre, P. and Kmiec, E.B.(1997) Isolation of human and mouse genes based on homology to REC2, a recombinational repair gene from the fungus *Ustilago maydis*. *Proc. Natl Acad. Sci. USA*, 94, 7417-7422
- Riley, R. and Chapman, V. (1958) Genetic control of the cytologically diploid behavior of hexaploid wheat. *Nature* 182, 713-715
- Roberts *et al.*, (1999) Induction and Characterization of *Ph1* Wheat Mutants *Genetics* 153, 1909–1918
- Rockmill B., Engebrecht J., Scherthan H., Loidl J., Roeder G.S. (1995) The yeast *MER2* gene is required for chromosome synapsis and the initiation of meiotic recombination. *Genetics* 141, 49–59
- Roder MS, Korzun V, Wandehake K, Planschke J, Tixier MH, Leroy P, Ganai MW (1998) A microsatellite map of wheat. *Genetics* 149, 2007-2023
- Roeder, G.S (1995) Sex and the single cell: meiosis in yeast. *Proc. Natl. Acad. Sci. USA*. 92, 10450-10456
- Romanienko, P.J. & Camerini-Otero, R.D, (1999) Cloning, Characterization, and Localization of Mouse and Human SPO11. *Genomics*. 61, 156-169
- Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B. 2003. An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Molecular Biology* 53, 247-259
- Rozwadowski, K.L. and Lydiate, D.J. Her Majesty the Queen in Right of Canada as represented by the Minister of Agriculture and Agrifood Canada. Modulation of meiotic recombination, WO/2002/022811S.L.
- Lusetti and M.M. Cox (2002) The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu. Rev. Biochem.* 71
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4, 406–425

- Sakamoto, S. (1973) Patterns of cytogenetic differentiation in the tribe Triticeae. *Seiken. Ziho* 41, 155-168
- Salmeron, J.M., Oldroyd, G.E.D., Rommens, C.M.T., Scofield, S.R., Kim, H., Lavelle, D.T., Dahlbeck, D. and Staskawicz, B.J. (1996) Tomato prf is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the Pto Kinase gene cluster. *Cell* 86, 123-133
- Sato S, Hotta Y, Tabata S. (1995) Structural analysis of a *RecA*-like gene in the genome of *Arabidopsis thaliana*. *DNA Research* 2, 89–93
- Schena Mark, Dari Shalon, Ronald W. Davis, Patrick O. Brown (1995) Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA. Microarray *Science* 270, 467 – 470
- Scherthan, H., Eils, R., Trelles-Sticken, E., Dietzel, S., Cremer, T., Walt, H. and Jauch, A. (1998) Aspects of three-dimensional chromosome reorganization during the onset of human male meiotic prophase. *J. Cell Sci.* 111, 2337-2351
- Scherthan, H., Weich, S., Schwegler, H., Härle, M., Heyting, C. and Cremer, T. (1996) Centromere and telomere movements during early meiotic prophase of mouse and man are associated with the onset of chromosome pairing. *J. Cell Biol.* 134, 109-1125
- Schild, D., Lio, Y.C., Collins, D.W., Tsomondo, T. and Chen, D.J. (2000) Evidence for simultaneous protein interactions between human Rad51 paralogs. *J. Biol. Chem.* 275, 16443–16449
- Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU (2005) A gene expression map of *Arabidopsis thaliana* development. *Nat Genet* 37, 501–506
- Schwacha A and Kleckner N., (1997) Interhomolog bias during meiotic recombination: meiotic functions promote a highly differentiated interhomolog only pathway. *Cell* 90, 1123–1135
- Schwacha and N. Kleckner, (1995) Identification of double Holliday junctions as intermediates in meiotic recombination *Cell* 83, 783–791

- Schwarzacher, T. (1997) Three stages of meiotic homologous chromosome pairing in wheat: cognition, alignment and synapsis. *Sexual Plant Reproduction* 10, 324-331
- Sears, E. R., (1977) An induced mutant with homoeologous pairing in common wheat. *Can. J. Genet. Cytol.* 19, 585-593
- Sears, E.R (1954) The aneuploids of common wheat. *Research Bulletin*, University of Missouri Agricultural Experiment Station, 1-58
- Sears, E.R. (1982) A wheat mutation conditioning an intermediate level of homoeologous chromosome pairing. *Can. J. Genet. Cytol.* 24, 715-719
- Sears, E.R. and Okamoto, M. (1958) Intergenomic chromosome relationships in hexaploid wheat, In *Proceedings of the Tenth International Congress of Genetics*, Montreal, Canada, pp. 258-259
- Sears, E.R., (1966) Nullisomic-tetrasomic combinations in hexaploid wheat, pp.29-45 in *Chromosome Manipulation and plant Genetics*, edited by R. RILEY and K.R.LEWIS. Oliver & Boyd, Edinburgh.
- Sears, E.R.and M.S.Sears, (1978). The telocentric chromosomes of common wheat, pp. 389-407 in *proceedings of the 5<sup>th</sup> International Wheat Genetics and Plant Breeding*. New Delhi.
- Seo, S., Maeda, T. and Hiratsuka, K. (2007) Tissue-specific and DNA damage-responsive expression of the Arabidopsis *RAD51* gene promoter in transgenic Arabidopsis and tobacco. *Plant Biotechnology* 24, 321-329
- Sessions, A.; Burke, E.; Presting, G.; Aux, G.; Mcelver, J.; Patton, D.; Dietrich, B.; Ho, P.; Backwaden, J.; KO, C.; Clarke, J.D.; Cotton, D.; Bullis, D.; Snell, J.; Miguel, T.; Hutchinson, D.; Kimmerly, B.; Mitzel, T.; Katagiri, F.; Glazebrook, J.; Law, M. and Goff, S.A (2002). A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell* 14, 2985-2994
- Shaked, H., C. Melamed-Bessudo and A. A. Levy, (2005) High-frequency gene targeting in Arabidopsis plants expressing the yeast RAD54 gene. *Proc. Natl. Acad. Sci. USA* 102, 12265-12269



- Shaked, H., Kashkush, K., Ozkan, H., Feldman, M., and Levy, A.A. (2001) Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* 13, 1749–1759
- Shao XG, Grishin NV (2000) Common fold in helix-hairpin-helix proteins. *Nucleic Acids Res.* 28:2643-2650
- Shapiro, G. I. (2006) Cyclin-Dependent Kinase Pathways As Targets for Cancer Treatment. *JCO* 24, 1770-1783
- Shimazu, J., Matsukura, C., Senda, M., Ishikawa, R., Akada, S., Harada, T., Tabata, S. and Niizeki, M. (2001) Characterization of *DMC1* homologue, *RiLIM15*, in meiotic panicles, mitotic cultured cells and mature leaves of rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 102, 1159-1163
- Shinohara A, Ogawa H, Matsuda Y, Ushio N, Ikeo K, Ogawa T. (1993) Cloning of human, mouse and fission yeast recombination genes homologous to *RAD51* and *RecA*. *Nature Genetics* 4, 239–243
- Shinohara A. and Ogawa, T. (1998) Stimulation by Rad52 of yeast Rad51-mediated recombination. *Nature* 391, 404–407
- Shinohara, A., Ogawa, H. and Ogawa, T. (1992) Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. *Cell* 69, 457-470
- Shitsukawa N, Ikari C, Mitsuya T, Sakiyama T, Ishikawa A, Takumi S, Murai K (2007b). Wheat *SOC1* functions independently of *WAP1/VRN1*, an integrator of vernalization and photoperiod flowering promotion pathways. *Physiol. Plant.* 130, 627–636
- Shitsukawa N, Ikari C, Shimada S, *et al.* (2007a) The einkorn wheat (*Triticum monococcum*) mutant, *maintained vegetative phase*, is caused by a deletion in the *VRN1* gene. *Genes Genet. Syst.* 82, 167–170
- Shitsukawa N, Tahira C, Kassai K, Hirabayashi C, Shimizu T, Takumi S, Mochida K, Kawaura K, Ogihara Y, Murai K (2007) Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. *Plant Cell* 19, 1723–1737

- Shu Z, Smith S, Wang L, Rice MC, Kmiec EB (1999) Disruption of muREC2/RAD51L1 in mice results in early embryonic lethality which can be partially rescued in a p53(-/-) background. *Mol Cell Biol* 19, 8686–8693
- Siaud N., Eloïse Dray, Isabelle Gy, Emmanuelle Gérard, Najat Takvorian and Marie-Pascale Doutriaux, (2004) Brca2 is involved in meiosis in *Arabidopsis thaliana* as suggested by its interaction with Dmcl, *EMBO J.* 23, 1392–1401
- SIFT: predicting amino acid changes that affect protein function, *Nucleic Acids Research*, 2003, Vol. 31, No. 13 3812-3814
- Sigurdsson S, Trujillo K, Song BW, Stratton S, Sung P (2001) Basis for avid homologous DNA strand exchange by human RAD51 and RPA. *J Biol Chem.* 276, 8798–8806
- Sigurdsson, S., Van Komen, S., Bussen, W., Schild, D., Albala, J.S. and Sung, P. (2001) Mediator function of the human Rad51B-Rad51C complex in Rad51/RPA-catalyzed DNA strand exchange. *Genes Dev.* 15, 3308–3318
- Slade A. J, Susan I Fuerstenberg, Dayna Loeffler, Micheal N steine and Daniel Facciottii, (2004) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nature Biotechnology* 23, 75 - 81
- Smith and Nicolas, (1998) K.H. Smith and A. Nicolas, Recombination at work for meiosis, *Curr. Opin. Genet. Dev.* 8, 200–211
- Solinger, J.A. Solinger, G. Lutz, T. Sugiyama, S.C. Kowalczykowski and W.D. Heyer, (2001) Rad54 protein stimulates heteroduplex DNA formation in the synaptic phase of DNA strand exchange via specific interactions with the presynaptic Rad51 nucleoprotein filament, *J. Mol. Biol.* 307, 1207–1221
- Soltis D. P. Soltis (1993) Molecular data and the dynamic nature of polyploidy. *Critical Reviews in the Plant Sciences* 12
- Soltis, D.E., P.S. Soltis, and J.A. Tate. (2004b) Advances in the study of polyploidy since plant speciation *New Phyto.* 161,173–191

- Soltis, D.E., P.S. Soltis, J.C. Pires, A. Kovarik, J.A. Tate, and E. Mavrodiev. (2004a) Recent and recurrent polyploidy in *Tragopogon* (Asteraceae): Cytogenetic, genomic and genetic comparisons. *Biol. J. Linn. Soc.* 82,485–501
- Stacey NJ, Kuromori T, Azumi Y, Roberts G, Breuer C, Wada T, Maxwell A, Roberts K, Sugimoto-Shirasu K (2006) *Arabidopsis* SPO11-2 functions with SPO11-1 in meiotic recombination. *Plant J* 48, 206–216
- Stadler., J. (1929). Chromosome number and mutation rate in *Avena* and *Triticum*. - *Proc. Not. Acad. Sci.* 15, 876-881
- Stadler., R. and A. M. Towe (1971) Evidence for meiotic recombination in *Ascobolus* involving only one member of a tetrad. *Genetics* 68, 401-413
- Stahl FW (1979) Special sites in generalized recombination. *Annu Rev Genet.*13, 7–24.
- Stassen N.Y., J.M. Logsdon Jr, G.J. Vora, H.H. Offenberger, J.D. Palmer and M.E. Zolan (1997) Isolation and characterization of rad51 orthologs from *Coprinus cinereus* and *Lycopersicon esculentum*, and phylogenetic analysis of eukaryotic *RecA* homologs. *Current Genetics* 31, 144–157
- Stebbins G. L. (1971) Chromosome evolution in higher plants. Addison-Wesley, Reading, Massachusetts, USA
- Stephens, S.G. (1951) Possible significance of duplication in evolution. *Adv. Genet.* 4, 247–265
- Stewart CN, Jr (2006) Go with the glow: fluorescent proteins to light transgenic organisms *Trends Biotechnol* 24: 155–162
- Story RM, Weber IT, Steitz TA (1992) The structure of the *E.coli* recA protein monomer and polymer. *Nature* 355, 318-324
- Sun, T.P., Goodman, H.M .and Ausubel, F.M. (1992) Cloning the *Arabidopsis* GA1 locus by genomic subtraction. *Plant Cell* 4, 119-128
- Sundaresan, V. (1996) Horizontal spread of transposon mutagenesis: new uses of ole elements. *Trends in Plant Sciences* 1, 184-191

- Sundaresan, V.; Springer, P.; Volpe, T.; Haward, S.; Jones, J.D.G.; Dean, C.; Ma, H. and Martienssen R, (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes and Development* 9, 1797-1810
- Sung (1997a) Function of yeast Rad52 protein as a mediator between replication protein-A and the Rad51 recombinase. *J Biol Chem.*272, 28194–28197
- Sung (1997b) Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein-A to promote DNA strand exchange by Rad51 recombinase. *Genes & Dev.*11, 1111–1121
- Sung P.(1994) Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast Rad51 protein. *Science* 265, 1241–1243
- Sung, P., Klein, H (2006) Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nature Reviews Molecular Cell Biology* 7, 739–750
- Symington, L. S., (2002) Role of *RAD52* epistasis group genes in homologous recombination and double-strand break repair. *Microbiol. Mol. Biol. Rev.* 66, 630–670
- Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW (1983) The double-strand-break repair model for recombination. *Cell* 33, 25–35
- Takata M, Sasaki MS, Tachiiri S, Fukushima T, Sonoda E, Schild D, Thompson LH, Takeda S (2001) Chromosome instability and defective recombinational repair in knockout mutants of the five Rad51 paralogs. *Mol Cell Biol* 21, 2858–2866
- Talbert LE, Smith LY, Blake MK (1998) More than one origin of hexaploid wheat is indicated by sequence comparison of low-copy DNA. *Genome* 41, 402–407
- Tambini,C.E., George,A.M., Rommens,J.M., Tsui,L.C., Scherer,S.W. and Thacker, J. (1997) The XRCC2 DNA repair gene: identification of a positional candidate. *Genomics*, 41, 84–92

- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 24, 1596–1599
- Tate, J.A., *et al.* (2006) Evolution and expression of homeologous loci in *Tragopogon miscellus* (Asteraceae), a recent and reciprocally formed allopolyploid. *Genetics* 173, 1599–1611
- Tebbs, R.S., Zhao, Y, M Tucker, J.D., Scheerer, J.B., Siciliano, M.J., Hwang, M., Liu, N, M Legerski, R.J. and Thompson, L.H. (1995) Correction of chromosomal instability and sensitivity to diverse mutagens by a cloned cDNA of the XRCC3 DNA repair gene. *Proc. Natl. Acad. Sci. USA* 92, 6354–6358
- Terada, R., Urawa, H., Inagaki, Y., Tsugane, K., and Iikda, S. (2002) Efficient gene targeting by homologous recombination in rice. *Nat. Biotechnol.* 20, 1030–1034
- Terasawa, M., Shinohara, A., Hotta, Y., Ogawa, H., and Ogawa, T. (1995) Localization of RecA-like protein in chromosomes of the lily at various meiotic stages. *Genes Dev.* 9, 925–934
- Thacker, J. (1999) A surfeit of RAD51-like genes? *Trends Genet* 15, 166–168
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25, 4876–4882
- Thompson, L.H. and Schild, D (2001) Homologous recombinational repair of DNA ensures mammalian chromosome stability. *Mutat. Res.* 477, 131–153
- Toshiyuki Kobayashi, Etsuko Kobayashi, Shusei Sato, Yasuo Hotta , Nobuyuki Miyajima, Ayako Tanaka, and Satoshi Tabata (1994) Characterization of cDNAs Induced in Meiotic Prophase in Lily Microsporocytes *DNA Res* 1, 15–26
- Trujillo KM, Yuan SSF, Lee and Sung P (1998) Nuclease activities in a complex of human recombination and repair factors Rad50, Mre11 and P<sup>95</sup> *J.Bio.Chem.* 273, 21447–21450

- Upadhyia, M.D and Swaminathan, M.S. (1967) Mechanism regulating chromosome pairing in Triticum, *Biol. Zentralbl. Suppl* 86, 239-255
- Urban, C., Smith, K.N. and Beier, H. (1996) Nucleotide sequences of nuclear tRNA(Cys) genes from Nicotiana and Arabidopsis and expression in HeLa cell extract. *Plant Mol. Biol.* 32, 549–552
- Usui, T., T. Ohta, H. Oshium, J. Tomizawa, and H. Ogawa *et al.*, (1998) Complex formation and functional versatility of Mre11 of budding yeast in recombination. *Cell* 95,705-716
- Van Campenhout S, Gebeyaw Z, Volckaert G (2003) Conversion of two RFLP probes to orthologue-specific STS-PCR markers and illustration of their applicability for comparative analysis of genome variation and relatedness in polyploidy wheat species. *Proc Wheat Genet Symp* 10, 533-535
- Van der Krol, A. R., Mur, L. A., Beld, M., Mol, J. N., and Stuitje, A. R. (1990). Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* 2, 291-299
- Vershon AK, Hollingsworth NM, Johnson AD (1992). Meiotic induction of the yeast HOP1 gene is controlled by positive and negative regulatory sites. *Mol Cell Biol* 12, 3706–3714
- Villeneuve AM, Hillers KJ.2001. Whence meiosis? *Cell* 106, 647-50
- Waines JG, Barnhart D (1992) Biosystematic research in *Aegilops* and *Triticum*. In: Proceedings of the 1st International Triticeae Symposium—Seberg O, Lundquist A, eds. Helsingborg, Sweden, Hereditas Offprint. 207–212
- Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) Distantly related equences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO Journal* 1, 945–951
- Wall, A.M., Riley, R., and Chapman, V (1971) Wheat mutants permitting homoeologous meiotic chromosomes pairing. *Genet. Res.* 18, 311–328
- Wang, J., *et al.* (2004) Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics* 167, 1961–1973

- Wang, J., *et al.* (2006b) Genome wide non-additive gene regulation in *Arabidopsis* allotetraploids. *Genetics* 172, 507–517
- Wang, J., L. Tian, H.-S. Lee, and Z.J. Chen. (2006a) Non-additive regulation of *FRI* and *FLC* loci mediates flowering-time variation in *Arabidopsis* allopolyploids. *Genetics* 173, 965–974
- Weiner, B. M. and Kleckner, N. (1994) Chromosome pairing via multiple interstitial interactions before and during meiosis in yeast. *Cell*. 77, 977–991
- Wendel, J., and J. Doyle (2005) Polyploidy and evolution in plants. P. 97–117. *In* R.J. Henry (ed.) Plant diversity and evolution: Genotypic and pheontypic variation in higher plants. CAB International.
- Wendel, J.F. (2000) Genome evolution in polyploids. *Plant Mol. Biol* 42, 225–249
- West, S.C. (2003). Molecular views of recombination and its control. *Nat. Rev. Mol. Cell Biol.* 4, 435–445
- Westergaard, M. and D. Wettstein (1972) The synaptonemal complex. *Ann.Rev.Genet* 6, 71–110
- Wiese, C., Collins, D.W., Albala, J.S., Thompson, L.H., Kronenberg, A. and Schild, D. (2002) Interactions involving the Rad51 paralogs Rad51C and XRCC3 in human cells. *Nucleic Acids Res.* 30, 1001–1008
- Wijeratne Asela J., Changbin Chen, Wei Zhang, Ljudmilla Timofejeva, and Hong Ma, (2006) The *Arabidopsis thaliana* *PARTING DANCERS* gene encoding a novel protein is required for normal meiotic homologous recombination, *Mol. Biol. Cell*, 17 1331–1343
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18, 6531–6535
- Winkler, H. (1916). Über die experimentelle Erzeugung von Pflanzen mit abweichenden Chromosomenzahlen. *Z. Bot.* 8, 417–531



- Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, *et al.* (2007) An “Electronic Fluorescent Pictograph” Browser for Exploring and Analyzing Large-Scale Biological Data Sets. *PLoS ONE* 2(8)
- Wittich, P.E. *et al.* Keygene N.V. Homologous recombination in plants, WO/2007/030014
- Wojciech P. Pawlowski, Inna N. Golubovskaya, Ljudmilla Timofejeva, Robert B. Meeley, William F. Sheridan, and W. Zacheus Cande (2004) Coordination of Meiotic Recombination, Pairing, and Synapsis by PHS1 *Science* 303, 89
- Wolfe K. H. (2001) Yesterday's polyploids and the mystery of diploidization. *Nat. Rev. Genet.* 2, 333–341
- Wu, L.M., Z.F. Ni, F.R. Meng, Z. Lin, and Q.X. Sun (2003) Cloning and characterization of leaf cDNAs that are differentially expressed between wheat hybrids and their parents. *Mol. Genet. Gen.* 270, 281–286
- Yazaki, J., Kojima, K., Suzuki, K., Kishimoto, N. and Kikuchi, S. (2004) The Rice PIPELINE: a unification tool for plant functional genomics. *Nucleic Acids Research* 32, 383-387
- Yoshida, K., Kondoh, G., Matsuda, Y., Habu, T., Nishimune, Y., and Morita, T. (1998). The mouse RecA-like gene Dmc1 is required for homologous chromosome synapsis during meiosis. *Mol. Cell* 1, 707-718
- Zadok, J.C., Chang, T.T., Konzak, (1974) A decimal code for the growth stages of cereals. *Eucarpia Bull* 7, 19
- Zhang XP, Lee KI, Solinger JA, Kiianitsa K, Heyer WD (2005) Gly-103 in the N-terminal domain of *Saccharomyces cerevisiae* Rad51 protein is critical for DNA binding. *J Biol Chem* 280, 26303-11
- Zimmermann Philip, Matthias Hirsch-Hoffmann, Lars Hennig, and Wilhelm Gruissem (2004) GENEVESTIGATOR Arabidopsis Microarray Database and Analysis Toolbox. *Plant Physiol* 136, 2621-2632